

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2003, 12:39:57 ; Search time 21 Seconds

(without alignments)
2238.559 Million cell updates/sec

Title: US-09-455-978b-2

Perfect score: 2394

Sequence: 1 MSNDNDLTVTADVNGIDGH.....ATDQVRYVEEVRVYKLS 489

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: PIR73:*
2: PIR1:*
3: PIR2:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2394	100.0	489	2 T44978	transducer protein
2	2360	98.6	489	2 E84304	Htr10 transducer
3	643	26.9	481	2 A84294	Htr9 transducer
4	609.5	25.5	482	2 T44973	transducer protein
5	597	24.9	633	2 H84336	Htr3 transducer
6	597	24.9	805	2 T48840	transducer protein
7	580	24.2	790	2 B84238	Htr18 transducer
8	578	24.1	778	2 F84237	Htr4 transducer
9	573	23.9	778	2 T48897	transducer protein
10	560.5	23.4	452	2 T44849	transducer protein
11	555	23.2	810	2 F84327	Htr5 transducer
12	554	23.1	810	2 T46810	halobacterial tran
13	540.5	22.6	777	2 T44597	transducer protein
14	537.5	22.5	534	2 S55299	sensory rhodopsin
15	535.5	22.4	636	2 A84252	Htr15 transducer
16	529	22.1	804	2 T44606	transducer protein
17	524	21.9	451	2 T44964	transducer protein
18	523.5	21.9	544	2 T44938	transducer protein
19	523.5	21.9	643	2 H84305	Htr8 transducer
20	521.5	21.8	628	2 F84219	Htr6 transducer
21	521	21.8	545	2 E84327	Htr7 transducer
22	521	21.8	545	2 T46811	halobacterial tran
23	509.5	21.3	789	2 E84236	Htr6 transducer
24	507.5	21.2	788	2 T44262	transducer protein
25	506	21.1	420	2 C84298	Htr12 transducer
26	505	21.1	642	1 T44253	transducer protein
27	502.5	21.0	773	2 T44989	Htr14 transducer
28	499.5	20.9	627	2 F84194	transducer protein
29	497.5	20.8	419	2 T44276	transducer protein

REST AVAILABLE COPY

30	497	20.8	536	2 D84325	Htr17 transducer
31	475.5	19.9	767	2 C87302	methy1-accepting c
32	464	19.4	534	2 A84328	Htr2 transducer
33	463.5	19.4	536	1 A47190	transducer protein
34	463.5	19.4	536	2 E84318	Htr1 transducer
35	461	19.3	765	1 T44946	transducer protein
36	452.5	18.9	423	2 H84257	Htr3 transducer
37	439.5	18.4	423	2 T44258	transducer protein
38	419.5	17.5	559	2 G84132	methy1-accepting c
39	415	17.3	555	2 D87536	methy1-accepting c
40	413	17.3	439	2 A85713	methy1-accepting c
41	412.5	17.2	432	2 C69832	chemotactic transd
42	409.5	17.1	632	2 H83106	methy1-accepting c
43	408	17.0	499	2 A97485	methy1-accepting c
44	408	17.0	499	2 A82703	methy1-accepting c
45	405.5	16.9	579	2 D84137	methy1-accepting c

ALIGNMENTS

RESULT 1	T44978	transducer protein hemA [validated] - Halobacterium salinarum
N:Alternate names: methy1-accepting taxis protein htb; transducer protein htb; transd		
C:Species: Halobacterium salinarum		
C:Date: 21-Jan-2000 #sequence: revision 21-Jan-2000 #text: change 15-Sep-2000		
C:Accession: T44978		
R: Zhang, W.; Brooun, A.; McCandless, J.; Banda, P.; Alam, M.		
Proc. Natl. Acad. Sci. U.S.A. 93, 4649-4654, 1996		
A:Title: Signal transduction in the archaeon Halobacterium salinarum is processed th		
A:Reference number: 222804; MUID:96209786; PMID:8643458		
A:Accession: T44978		
A:Status: preliminary; translated from GB/EMBL/DBJ		
A:Molecule type: DNA		
A:Residues: 1-489 <ZHA>		
A:Cross-references: EMBL:U75436; NID:g1654420; PIDN:AA817881.1; PID:g1654421		
A:Experimental source: strain Flx15		
A:Note: the source is designated as Halobacterium salinarum		
C:Genetics:		
A:Gene: hemA; htp15; htb		
C:Function:		
A:Description: involved in aerotactic signal transduction; involved in oxygen sensing		
C:Superfamily: Halobacterium salinarum transducer protein htr1		
C:Keywords: heme; methylated amino acid; signal transduction		
Query Match	100.0%; Score 2194; DB 2; Length 489;	
Best local similarity	100.0%; Pred. No. 8.4e-98;	
Matches	489; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MSNDNDLTVTADYRNGIDGHADLRIGLDEAEIAMLSTFGIDDDPMALAAEQPFERT 60	
DB	1 MSNDNDLTVTADYRNGIDGHADLRIGLDEAEIAMLSTFGIDDDPMALAAEQPFERT 60	
QY	61 AADLVNDFPDHLESYRTODLFANSRTVEQLKETQAEVLGGRGEYDEYVAOARRIG 120	
DB	61 AADLVNDFPDHLESYRTODLFANSRTVEQLKETQAEVLGGRGEYDEYVAOARRIG 120	
QY	121 KIHVDVGLGPDVYLGAATRYTGLDLADVDVADAGEEAAAADVLAFLPMLKLTF 180	
DB	121 KIHVDVGLGPDVYLGAATRYTGLDLADVDVADAGEEAAAADVLAFLPMLKLTF 180	
QY	181 DOOIAMDTYIDSYAORLHDEIDSRGLANAVATHEAPLSSLEATSDVAERTDTRART 240	
DB	181 DOOIAMDTYIDSYAORLHDEIDSRGLANAVATHEAPLSSLEATSDVAERTDTRART 240	
QY	241 DDVDMAVVSREISSVSAAVEEASTADVRRTSSDAEALAOGEAAADALATMTDID 300	
DB	241 DDVDMAVVSREISSVSAAVEEASTADVRRTSSDAEALAOGEAAADALATMTDID 300	
QY	301 EATDGYTAGVEQLGERAAVESYTGVIDIAETNNLALNASTEARAGEAGGFVAVD 360	
DB	301 EATDGYTAGVEQLGERAAVESYTGVIDIAETNNLALNASTEARAGEAGGFVAVD 360	

```

QY 361 EVKALAEESRQSTREVELVEQMAETEETVDQLDEVNQRIGGVERVEEAMETLQETID 420
DB 361 EVKALAEESRQSTREVELVEQMAETEETVDQLDEVNQRIGGVERVEEAMETLQETID 420
QY 421 AVEADAASGMOEYSTATDEQAVSTEEVAEMVGVDDRRAGEIAALADLDIADATDOQVTRVEE 480
DB 421 AVEADAASGMOEYSTATDEQAVSTEEVAEMVGVDDRRAGEIAALADLDIADATDOQVTRVEE 480
QY 481 VRETVGKLS 489
DB 481 VRETVGKLS 489

RESULT 2
Hit10 transducer (imported) - Halobacterium sp. NRC-1
E84304
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: E84304
R:Ng, W.V.; Kennedy, S.P.; Mahafiras, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Laaky, S.;
Jung, K.H.; Alam, M.; Freltas, T.;
Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Edhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: E84304
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-489 <STO>
A:Cross-references: GB:AE004437; NID:g10580997; PIDN:AG19801.1; GSPDB:GN00138
C:Genetics:
A:Gene: htr10
C:superfamily: Halobacterium salinarum transducer protein htr1

Query Match 98.6%; Score 2360; DB 2; Length 489;
Best Local Similarity 99.0%; Pred. No. 2.6e-96;
Matches 484; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MSNDNPTLVADYRNGIDGHALDRIGLDEAEIAMLRLSFTGIDDDTMAALAEQPLFEAT 60
DB 1 MSNDNPTLVADYRNGIDGHALDRIGLDEAEIAMLRLSFTGIDDDTMAALAEQPLFEAT 60
QY 61 ADALVTDFYDHLESYERTDQLFANSTKTYEQLEKQAEYLLGLRGEDYTEVAQRRIG 120
DB 61 ADALVTDFYDHLESYERTDQLFANSTKTYEQLEKQAEYLLGLRGEDYTEVAQRRIG 120
QY 121 KIRIDVGLGPDYLGATRYTGTGLDALDDVYADRGEAAAVDELVARFLPMLKLTF 180
DB 121 KIRIDVGLGPDYLGATRYTGTGLDALDDVYADRGEAAAVDELVARFLPMLKLTF 180
QY 121 KIRIDVGLGPDYLGATRYTGTGLDALDDVYADRGEAAAVDELVARFLPMLKLTF 180
DB 121 KIRIDVGLGPDYLGATRYTGTGLDALDDVYADRGEAAAVDELVARFLPMLKLTF 180
QY 181 DOOIADMTYIDSYAORLHDEIDSRQELANAVATHVEAPLSSLETSODVNERDTMRAPT 240
DB 181 DOOIADMTYIDSYAORLHDEIDSRQELANAVATHVEAPLSSLETSODVNERDTMRAPT 240
QY 241 DDQVDRMADYSRRISSVSASVEEYASTADVRRTSEDAELAQOGEAADALATMTDID 300
DB 241 DDQVDRMADYSRRISSVSASVEEYASTADVRRTSEDAELAQOGEAADALATMTDID 300
QY 241 DDQVDRMADYSRRISSVSASVEEYASTADVRRTSEDAELAQOGEAADALATMTDID 300
DB 241 DDQVDRMADYSRRISSVSASVEEYASTADVRRTSEDAELAQOGEAADALATMTDID 300
QY 301 EATDGTAGYEQLCERADYESTGYIDTAEQTNMLANASIEARAGAGGFAVVAD 360
DB 301 EATDGTAGYEQLCERADYESTGYIDTAEQTNMLANASIEARAGAGGFAVVAD 360
QY 301 EATDGTAGYEQLCERADYESTGYIDTAEQTNMLANASIEARAGAGGFAVVAD 360
DB 301 EATDGTAGYEQLCERADYESTGYIDTAEQTNMLANASIEARAGAGGFAVVAD 360
QY 361 EVKALAEESRQSTREVELVEQMAETEETVDQLDEVNQRIGGVERVEEAMETLQETID 420
DB 361 EVKALAEESRQSTREVELVEQMAETEETVDQLDEVNQRIGGVERVEEAMETLQETID 420
QY 421 AVEADAASGMOEYSTATDEQAVSTEEVAEMVGVDDRRAGEIAALADLDIADATDOQVTRVEE 480
DB 421 AVEADAASGMOEYSTATDEQAVSTEEVAEMVGVDDRRAGEIAALADLDIADATDOQVTRVEE 480
QY 481 VRETVGKLS 489
DB 481 VRETVGKLS 489
```

```

DB 481 VRETVGKLS 489

RESULT 3
Hit9 transducer (imported) - Halobacterium sp. NRC-1
A84294
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 17-May-2002
C:Accession: A84294
R:Ng, W.V.; Kennedy, S.P.; Mahafiras, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Laaky, S.;
Jung, K.H.; Alam, M.; Freltas, T.;
Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Ja
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Edhardt, H.; Lowe, T.M.;
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: A84294
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-481 <STO>
A:Cross-references: GB:AE004437; NID:g10580901; PIDN:AG19717.1; GSPDB:GN00138
C:Genetics:
A:Gene: htr9
C:superfamily: Halobacterium salinarum transducer protein car

Query Match 26.9%; Score 643; DB 2; Length 481;
Best Local Similarity 37.6%; Pred. No. 2.2e-21;
Matches 182; Conservative 72; Mismatches 196; Indels 34; Gaps 10;

QY 15 NGIDGHALDRIGLDEAEIAMLRLSFTGIDDDTMAALAEQPLFEATADAL-VTDFYDHE 73
DB 15 NGIDGHALDRIGLDEAEIAMLRLSFTGIDDDTMAALAEQPLFEATADAL-VTDFYDHE 73
QY 19 NTLDPVPAFA--VDADGAVVAM-----DDQIAL--LETADEDAIGVTDIGERLN 63
DB 19 NTLDPVPAFA--VDADGAVVAM-----DDQIAL--LETADEDAIGVTDIGERLN 63
QY 74 SYERTDQLFANSTKTYEQLEKQAEYLLGLRGEDYTEVAQRRIGKIRIDVGLGPDY 133
DB 74 SYERTDQLFANSTKTYEQLEKQAEYLLGLRGEDYTEVAQRRIGKIRIDVGLGPDY 133
QY 64 D----DGRALANKVADTPIDAHNEY---DGLADESTALLTGVDYEDTTVAGNTDLW 116
DB 64 D----DGRALANKVADTPIDAHNEY---DGLADESTALLTGVDYEDTTVAGNTDLW 116
QY 134 LGATRYTGTGLDALDDVYADRGEAAAVDELVARFLPMLKL-----TFDQOI-- 184
DB 134 LGATRYTGTGLDALDDVYADRGEAAAVDELVARFLPMLKL-----TFDQOI-- 184
QY 117 FINTPYVHGEFFRGVI-EIVQDR-SSARQOSLGALFGLVPTLDAYDAGRGFDATVDIA 174
DB 117 FINTPYVHGEFFRGVI-EIVQDR-SSARQOSLGALFGLVPTLDAYDAGRGFDATVDIA 174
QY 185 AMDTYIDSYAORLHDEIDSRQELANAVATHVEAPLSSLETSODVNERDTMRAPDDQY 244
DB 185 AMDTYIDSYAORLHDEIDSRQELANAVATHVEAPLSSLETSODVNERDTMRAPDDQY 244
QY 175 AEDTLLDDEYIQIRNLTEGDTLAAHTEVHNDVVERLEAASQAVSESSAEIDELSTRAQS 234
DB 175 AEDTLLDDEYIQIRNLTEGDTLAAHTEVHNDVVERLEAASQAVSESSAEIDELSTRAQS 234
QY 245 DRMADYSRRISSVSASVEEYASTADVRRTSEDAELAQOGEAADALATMTDIEAPT*304
DB 245 DRMADYSRRISSVSASVEEYASTADVRRTSEDAELAQOGEAADALATMTDIEAPT*304
QY 235 TNVSTVATEVEYTSATVQELASTADEYVDTSATAERLADDGSAASDAADMAADVATAAD 294
DB 235 TNVSTVATEVEYTSATVQELASTADEYVDTSATAERLADDGSAASDAADMAADVATAAD 294
QY 305 GVTAGVEQLGERADYESTGYIDTAEQTNMLANASIEARAGAGGFAVVAD 364
DB 305 GVTAGVEQLGERADYESTGYIDTAEQTNMLANASIEARAGAGGFAVVAD 364
QY 295 SVTSDEVALQNRJEDIDEVVDVITGIAEQTNMLANASIEARAGAGGFAVVAD 354
DB 295 SVTSDEVALQNRJEDIDEVVDVITGIAEQTNMLANASIEARAGAGGFAVVAD 354
QY 365 LAESRQSTREVELVEQMAETEETVDQLDEVNQRIGGVERVEEAMETLQETID 424
DB 365 LAESRQSTREVELVEQMAETEETVDQLDEVNQRIGGVERVEEAMETLQETID 424
QY 355 LAEDAQSNMAGHISLVSEIQRDADYVTDLYTTDRIEDVAVQEDMAAFEEIYAVAVRA 414
DB 355 LAEDAQSNMAGHISLVSEIQRDADYVTDLYTTDRIEDVAVQEDMAAFEEIYAVAVRA 414
QY 425 AASGMOEYSTATDEQAVSTEEVAEMVGVDDRRAGEIAALADLDIADATDOQVTRVEE 484
DB 425 AASGMOEYSTATDEQAVSTEEVAEMVGVDDRRAGEIAALADLDIADATDOQVTRVEE 484
QY 415 TAEGIEVSDATNEQASAEETIAMVDETADLDITTAADIVSOTETASAMHLDDLS 474
DB 415 TAEGIEVSDATNEQASAEETIAMVDETADLDITTAADIVSOTETASAMHLDDLS 474
QY 485 VGRKL 488
DB 485 VGRKL 488
QY 475 VSEL 478
DB 475 VSEL 478

RESULT 4
transducer protein htr1 [similarity] - Halobacterium salinarum
N:Alternate names: methyl-accepting taxis protein htr; transducer protein htr; trans
C:Species: Halobacterium salinarum
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 02-Mar-2001
```

```

RESULT 5
HB4336
Htr3 transducer [Imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 17-May-2002
C:Accession: HB4336
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
: Leithauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocke, D.G.; Jablon
: Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li
A:Title: Genome Sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: HB4336

```

A:Residues: 1-805 <KOK>
A:Cross-references: EMBL:AJ245590; PID:CBH82572.1
A:Experimental source: strain S9
R:Zhang, W.; Brooun, A.; McCandless, J.; Banda, P.; Alm, M.
Proc. Natl. Acad. Sci. U.S.A. 93, 4649-4654, 1996
A>Title: Signal transduction in the archaeal Halobacterium salinarum is processed through a novel protein
A:Reference number: 222804; MUID:96209786; PMID:8643458
A:Accession: T44981
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 'M',13-128,'R',130-133,'S',135-200,'MP',203-241,'R',243-332,'MR',335-401.
<HHA>
A:Cross-references: EMBL:U75437; NID:91654422; PID:NAB17882.1; PID:91654423
A:Experimental source: strain Flx15 derivative of S9

F84327
Htr5 transducer [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: F84327
R:Ng, W.V.; Kennedy, S.P.; Mahapatra, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Laaky, S.; Leithauer, B.; Kellier, K.; Cruz, R.; Danison, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Jung, K.H.; Alam, M.; Freltas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; M0ID:20504483; PMID:11016950
A:Accession: F84327
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-810 <STR>
A:Cross-references: GB:AE004437; NID:g10581215; PIDN:AG19986.1; GSPDB:GN00138
C:Genetics:
A:Gene: htr5
C:Superfamily: Halobacterium salinarum transducer protein htrII

Query Match 23.2%; Score 555; DB 2; Length 810;
Best Local Similarity 32.1%; Pred. No. 2.8e-17;
Matches 142; Conservative 85; Mismatches 174; Indels 42; Gaps 8;

QY 84 NSTKVEQLKETQAEYLLGRCGEYDTEVAQR-----ARIGKHVYLGCPDYLLG 135
DB 346 NTAAGVQSLSAAAE-----IEAGNYDVVASSRDEIGQFASIGSMRDALVYQIDBAEA 401
QY 136 AYTRYTGLDALAD---DVVADRGEEAAAVDELVA-----RFLPKLL----- 178
DB 402 ARQATEAQDAAEERERADAREDAKADAEALAEQAERYSDVMAACADGDLT 461
QY 179 -----TFDQQA--MDTYIDSYAQRHLHEIDSRQELANAVT---HVEAPLSLEATS 226
DB 462 RRPADDTDNEAATAASNEMLAQWEHTIDI-QEFADAVATASEEAVGAADERAS 520
QY 227 QDVAERTDTRARTDQVDNADVSREISSVSAVEEASTADVRRTSSEDAALAOGE 286
DB 521 GQVSESVQETAGAACDEQRNMLDTVSGEMTDLSAIEEVAASDVASHSHQTAIARDGE 580
QY 287 AAADDALATMTDIDEATDGTAGVGEOLGERADVESYGVYIDIAQTNNLALNASTEA 346
DB 581 QTAEDAIERSLVQEAIDATVQVNEALDDQMAEISEIVLDSIAQTNNLALNANIEAA 640
QY 347 RAGEGEGFVAVVADYKALAESESTREVELEVMQATEETVQDLDEVNORIGGVE 406
DB 641 RAKSGDGFVAVVADYKALAESESTREVELEVMQATEETVQDLDEVNORIGGVE 406
QY 407 RVEAMETLOETIDAVEDAASGQEVSTATDEQAVSTEVEAVNDGVDRAGEITAAALDD 466
DB 701 AVEEVVDAFAVAVSDHADETGTQVEISDTTDDQASTEEAVSMTEEVADLSDTAGBAQS 760
QY 467 IADATDQVTRVEEVRVETGKLS 489
DB 761 VSAAEQAASMSSEISDSVESLS 783

RESULT 12

T46810
halobacterial transducer protein IV [imported] - Halobacterium salinarum
C:Species: Halobacterium salinarum
C:Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 02-Sep-2000
C:Accession: T46810
R:Rudolph, J.; Nordmann, B.; Storch, K.F.; Grunberg, H.; Rodewald, K.; Oesterhelt, D.
FEBS Microbiol. Lett. 139, 161-168, 1996
A:Title: A family of halobacterial transducer proteins.
A:Reference number: Z24094; M0ID:96275896; PMID:8674984
A:Accession: T46810
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-810 <RUD>
A:Cross-references: EMBL:X95589; NID:g1435130; PIDN:CA644841.1; PID:g1435131

A:Experimental source: strain S9
C:Genetics:
A:Gene: htrIV
C:Superfamily: Halobacterium salinarum transducer protein htrII

Query Match 23.1%; Score 554; DB 2; Length 810;
Best Local Similarity 32.1%; Pred. No. 3.1e-17;
Matches 142; Conservative 85; Mismatches 174; Indels 42; Gaps 8;

QY 84 NSTKVEQLKETQAEYLLGRCGEYDTEVAQR-----ARIGKHVYLGCPDYLLG 135
DB 346 NTAAGVQSLSAAAE-----IEAGNYDVVASSRDEIGQFASIGSMRDALVYQIDBAEA 401
QY 136 AYTRYTGLDALAD---DVVADRGEEAAAVDELVA-----RFLPKLL----- 178
DB 402 ARQATEAQDAAEERERADAREDAKADAEALAEQAERYSDVMAACADGDLT 461
QY 179 -----TFDQQA--MDTYIDSYAQRHLHEIDSRQELANAVT---HVEAPLSLEATS 226
DB 462 RRPADDTDNEAATAASNEMLAQWEHTIDI-QEFADAVATASEEAVGAADERAS 520
QY 227 QDVAERTDTRARTDQVDNADVSREISSVSAVEEASTADVRRTSSEDAALAOGE 286
DB 521 GQVSESVQETAGAACDEQRNMLDTVSGEMTDLSAIEEVAASDVASHSHQTAIARDGE 580
QY 287 AAADDALATMTDIDEATDGTAGVGEOLGERADVESYGVYIDIAQTNNLALNASTEA 346
DB 581 QTAEDAIERSLVQEAIDATVQVNEALDDQMAEISEIVLDSIAQTNNLALNANIEAA 640
QY 347 RAGEGEGFVAVVADYKALAESESTREVELEVMQATEETVQDLDEVNORIGGVE 406
DB 641 RAKSGDGFVAVVADYKALAESESTREVELEVMQATEETVQDLDEVNORIGGVE 406
QY 407 RVEAMETLOETIDAVEDAASGQEVSTATDEQAVSTEVEAVNDGVDRAGEITAAALDD 466
DB 701 AVEEVVDAFAVAVSDHADETGTQVEISDTTDDQASTEEAVSMTEEVADLSDTAGBAQS 760
QY 467 IADATDQVTRVEEVRVETGKLS 489
DB 761 VSAAEQAASMSSEISDSVESLS 783

RESULT 13

T44597
transducer protein htrVI [similarity] - Halobacterium salinarum
N:Alternate names: methyl-accepting transducer protein htd; transducer protein htr6
C:Species: Halobacterium salinarum
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 02-Sep-2000
C:Accession: T44597
R:Zhang, W.; Brooun, A.; McCandless, J.; Bando, P.; Alam, M.
Proc. Natl. Acad. Sci. U.S.A. 93, 4649-4654, 1996
A:Title: Signal transduction in the archaeon Halobacterium salinarum is processed th
C:Keywords: methylated amino acid; signal transduction; transmembrane protein
F:494-746/Region: MCP signalling domain similarity
A:Note: the source is designated as Halobacterium salinarum
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-777 <ZHA>
A:Cross-references: EMBL:U75438; NID:g1654424; PIDN:AA817883.1; PID:g1654425
A:Experimental source: strain mfix15
A:Note: the source is designated as Halobacterium salinarum
C:Genetics:
A:Gene: htrVI; htr6; htd
A:Superfamily: Halobacterium salinarum transducer protein htrII
C:Keywords: methylated amino acid; signal transduction; transmembrane protein
F:494-746/Region: MCP signalling domain similarity

Query Match 22.6%; Score 540.5; DB 2; Length 777;
Best Local Similarity 32.4%; Pred. No. 1.1e-16;
Matches 141; Conservative 74; Mismatches 171; Indels 49; Gaps 6;

QY 88 TVEQLEKETAQAEYLLGRCGEYDTEVAQR--RIGKIHVYLGCPDYLLGAYTRYTGLLD 146
DB 320 TVRALNDLEAK-APALERGEYTDLDVAVRDEGLRLEAFASIRLDVAVORADSDANQOVD 378

Oy 147 ALA-----DDVADRGEEAAAANDVLARFLP 173
 Db 379 AEAARSEAAQAEBEAAQAEAAAREESDAGAPGDGRRGVGLGGDACVAAGDLTVR--- 435
 Oy 174 MLKLTTFPOOLA----MPTIYDSYAQRHLHDEIDSRQELANAVA---TVEAPLSLEATS 226
 Db 436 ----LDADVEGAAMADMARRAFENMAADMEATTAEFGARADEVATATASTASDSAAVEOTG 491
 Oy 227 ODVAERTDMRAITDDQVDNRADVSREISSVASAYEEVAASTADVVRTSEDAEALAOGE 286
 Db 492 RDVSTAVERIRIFRADQRQLQLEAVASEFTDEMKSATIEEVARVAGVAETTSQRAALGDDQG 551
 Oy 287 AAADALATMTDIDEATGCGVTAGVEQJLGERADVESYGVDDIAEQTNMLALNASTAA 346
 Db 552 AAAQAVAKOLEIEEDETOAAATVADLLEKKSEIETIAATIDAQTNMALNLNNIEAA 611
 Oy 347 RAGEAGEGFAVAVADVKALAERSREGOSTRVELVEYQOMQOETEETVDOQJDEVNQRIEGEVE 406
 Db 612 RADQGGDGFAVADDEVKDLADESKALAAEMQALVAEVRAGQETTSVAAMDRCIOERSDGE 671
 Oy 407 RVEEMETLOETTTDAVEDAASGMQEVSTATDEQAASVEEVAEMVDGVDRAGETIAALD 466
 Db 672 TVSETERSLSDIAGRARIAPEDTGCEISNAMDDQAAVSVDVTAVAGVALGGEATTEAES 731
 Oy 467 IADATDQGVRTVEEV 481
 Db 732 TRDPAAEQAATTLSDV 746

RESULT 14
 S55299
 sensory rhodopsin II transducer protein - Natronobacterium pharaonis
 C:Species: Natronobacterium pharaonis
 C:Date: 14-Oct-1995 #sequence revision 03-Nov-1995 #text change 02-Sep-2000
 C:Accession: S55299
 R:Seidel, R.; Scherf, B.; Gautel, M.; Kleine, K.; Oesterheld, D.; Engelhard, M.
 Proc. Natl. Acad. Sci. U.S.A. 92, 3036-3040, 1995
 A:title: The primary structure of sensory rhodopsin II: a member of an additional retinal
 A:Reference number: S55296; MUID:95224074; PMID:7708770
 A:Accession: S55299
 A>Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-534 <SEI>
 A:Cross-references: EMBL:Z35086; NID:g510868; PID:g510866
 A:Experimental source: strain SP1
 A>Note: It is uncertain whether Met-1 or Met-22 is the initiator
 C:Genetics:
 A:Gene: htrII
 C:Superfamily: Halobacterium salinarum transducer protein htrI
 C:Keywords: signal transduction; transmembrane protein
 F:22-42/Domain: transmembrane #status predicted <TM>
 F:59-79/Domain: transmembrane #status predicted <TM>

Query Match	22.5%	Score 537.5;	DB 2;	Length 534;
Best Local Similarity	31.2%	Pred. No. 9.8e-17;		
Matches 150;	Conservative 95;	Mismatches 177;	Indels 59;	Gaps 12;

```

Oy   16 GIDGHALADRIGLDER-----EIAMRLSFTGIDDDTMALAAEQPLFEATADALVTDYF-- 69
Db   72 GINTGLIVATLGGDPNAASLSTLAAKASRMG-DGDDLDEL-----ERRREDICIDLTYAA 123
Oy   70 -DHLESYERTQDLFANSRKTVBOLKEQAEYLGLGRGEYDPEYVARAGARKIHDIYGL 128
Db   124 PDEMQRVSTSLIEDKNKAREDAQKRRAE-----EINTELQAEKREFGEVMDCRAD 175
Oy   139 GPDYVLGAVTRYRYTLLDALDDVVADNGEEAAAAVDIELVARFLPMKLILTFDQAIADT 188
Db   176 GD-----FTORLDAETDN-----EAMQSIEG-----SFNEH--MDG 204
Oy   169 YIDTAQRRLHEDIIDSROELANNAVATHVEAPLSLEATSGODVAERTDMRAITDDQVDNR 248
Db   205 -LEALVGRTEREPADVSEDAEVRANAEE---SYMEA-SEDVNRAVONISDAAGDTEFTVO 259
  
```

QY	249	DVSRKISSVSA	VEEVA	STAD	DDV	VRT	SEDA	EAL	QOGEA	AD	ALAT	TT	ID	EA	TC	GYTA	308
QY	249	DVSRKISSVSA <td>VEEVA<td>STAD<td>DDV<td>VRT<td>SEDA<td>EAL<td>QOGEA<td>AD<td>ALAT<td>TT<td>ID<td>EA<td>TC<td>GYTA<td>308</td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	VEEVA <td>STAD<td>DDV<td>VRT<td>SEDA<td>EAL<td>QOGEA<td>AD<td>ALAT<td>TT<td>ID<td>EA<td>TC<td>GYTA<td>308</td></td></td></td></td></td></td></td></td></td></td></td></td></td>	STAD <td>DDV<td>VRT<td>SEDA<td>EAL<td>QOGEA<td>AD<td>ALAT<td>TT<td>ID<td>EA<td>TC<td>GYTA<td>308</td></td></td></td></td></td></td></td></td></td></td></td></td>	DDV <td>VRT<td>SEDA<td>EAL<td>QOGEA<td>AD<td>ALAT<td>TT<td>ID<td>EA<td>TC<td>GYTA<td>308</td></td></td></td></td></td></td></td></td></td></td></td>	VRT <td>SEDA<td>EAL<td>QOGEA<td>AD<td>ALAT<td>TT<td>ID<td>EA<td>TC<td>GYTA<td>308</td></td></td></td></td></td></td></td></td></td></td>	SEDA <td>EAL<td>QOGEA<td>AD<td>ALAT<td>TT<td>ID<td>EA<td>TC<td>GYTA<td>308</td></td></td></td></td></td></td></td></td></td>	EAL <td>QOGEA<td>AD<td>ALAT<td>TT<td>ID<td>EA<td>TC<td>GYTA<td>308</td></td></td></td></td></td></td></td></td>	QOGEA <td>AD<td>ALAT<td>TT<td>ID<td>EA<td>TC<td>GYTA<td>308</td></td></td></td></td></td></td></td>	AD <td>ALAT<td>TT<td>ID<td>EA<td>TC<td>GYTA<td>308</td></td></td></td></td></td></td>	ALAT <td>TT<td>ID<td>EA<td>TC<td>GYTA<td>308</td></td></td></td></td></td>	TT <td>ID<td>EA<td>TC<td>GYTA<td>308</td></td></td></td></td>	ID <td>EA<td>TC<td>GYTA<td>308</td></td></td></td>	EA <td>TC<td>GYTA<td>308</td></td></td>	TC <td>GYTA<td>308</td></td>	GYTA <td>308</td>	308
Db	260	Q:TALEMDV	VSAT <td>EEVA</td> <td>SADD</td> <td>IAKT</td> <td>AAQAE</td> <td>ET<td>EG</td><td>RET<td>ET<td>AT<td>EN<td>VE</td><td>S</td><td>RE</td><td>319</td></td></td></td></td></td>	EEVA	SADD	IAKT	AAQAE	ET <td>EG</td> <td>RET<td>ET<td>AT<td>EN<td>VE</td><td>S</td><td>RE</td><td>319</td></td></td></td></td>	EG	RET <td>ET<td>AT<td>EN<td>VE</td><td>S</td><td>RE</td><td>319</td></td></td></td>	ET <td>AT<td>EN<td>VE</td><td>S</td><td>RE</td><td>319</td></td></td>	AT <td>EN<td>VE</td><td>S</td><td>RE</td><td>319</td></td>	EN <td>VE</td> <td>S</td> <td>RE</td> <td>319</td>	VE	S	RE	319
QY	309	GVEO	GERA	DEV	ST	GV	TD	IA	EOT	N	MLA	NS	I	EA	R	AG	368
Db	320	SMEEL	NED	VE	RG	EV	SE	MI	AD	EA	Q	T	N	LA	NS	I	379
QY	365	SREO	STR	VEL	VE	Q	AE	T <td>EE</td> <td>T<td>VD</td><td>L</td><td>DE</td><td>V</td><td>N</td><td>RI</td><td>428</td></td>	EE	T <td>VD</td> <td>L</td> <td>DE</td> <td>V</td> <td>N</td> <td>RI</td> <td>428</td>	VD	L	DE	V	N	RI	428
Db	380	TKAAT	EE	IT	DL	IG	TV	Q	D	R	Q	T	T	VD	I	R	439
QY	429	MOE	VA	T	D	EA	Q	AV	ST <td>EE</td> <td>VA</td> <td>E</td> <td>MD</td> <td>V</td> <td>D</td> <td>R</td> <td>488</td>	EE	VA	E	MD	V	D	R	488
Db	440	IG	E	I	N	O	S	T	D	A	A	A	K	A	T	T	499
QY	489	S	489	S	489	S	489	S	489	S	489	S	489	S	489	S	489
Db	500	S	500	S	500	S	500	S	500	S	500	S	500	S	500	S	500
RESULT	15																
	A84252																
	Hrt15	transducer [imported] - Halobacterium sp. NRC-1															
	C:Species:	Halobacterium sp. NRC-1															
	C:Date:	02-Feb-2001															
	C:Accession:	A84252															
	R:Ng,	M.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, R.; Lettshauser, B.; Keller, K.; Cruz, R.; Denson, M.J.; Hough, D.W.; Maddocks, D.G.; Jung, K.H.; Alam, M.; Freitas, T.															
	Proc. Natl. Acad. Sci. U.S.A.	97, 12176-12181, 2000															
	A:Authors:	Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; A:Reference:															
	A:Reference:	Genome sequence of Halobacterium species NRC-1.															
	A:Accession:	A84252															
	A:Status:	preliminary															
	A:Molecule	type: DNA															
	A:Residues:	1-636 <STC>															
	A:Cross-references:	GB:AE004437; NID:g10580b13; PIDN:ANG19381.1; GSPDB:GN00138															
	A:Gene:	hrt15															
Query Match	22.4%	Score 535.5; DB 2; Length 636;															
Best Local Similarity	29.7%	Pred. No. 1.5e-16;															
Matches 149; Conservative	99;	Mismatch 191; Indels 63; Gaps 11;															
QY	17	IDGH	LAD	RIG	LDE	AE	I	AW	LS	FT	G	I	D	D	T	MA	67
Db	165	VDD	AL	L	LS	I	G	MP	----								

Thu Jan 2 14:11:53 2003

us-09-455-978b-2.rpr

Page 8

Db 491 AGEAGSGSFAVADVEKTELANETRHTERIRINGSISDVOQANNEVFLAVESHEQIRHAGE 550

Qy 408 VEEAENETJOTLTDVEAASGMOGVSATDEQASVTEVNAEMDGDVDRRGETAALDOI 467

Db 551 IDDLTALTEELATSVDEANCTITEVANAANDQASTVEDVLTITIEDVOOAAEENAAASDKI 610

Qy 468 ADATDOQVRYEEVREYTKUS 489

Db 611 VSAFOGOSTAVSOLSERVDKLT 632

```
Search completed: January 2, 2003, 12:43:10
Job time : 23 secs
```

• • •

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2003, 12:39:17 : Search time 35 Seconds

(Without alignments)
2878.773 Million cell updates/sec

Title: US-09-455-978b-2

Perfect score: 2394
Sequence: 1 MSNDNDTLVADVRNGIDGH.....ATDQVRYVEVRETVGKLS 489

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_oranelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protist:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	609.5	25.5	482	1 P71412	P71412 halobacteri
2	597	24.9	633	17 Q9HP10	Q9HP10 halobacteri
3	597	24.9	805	1 Q9P930	Q9P930 halobacteri
4	580	24.2	790	17 Q9HR88	Q9HR88 halobacteri
5	560.5	23.4	452	1 Q93775	Q93775 halobacteri
6	540.5	22.6	777	1 P71415	P71415 halobacteri
7	539	22.5	792	1 P71414	P71414 halobacteri
8	535.5	22.4	636	17 Q9H0X8	Q9H0X8 halobacteri
9	529	22.1	804	1 P71416	P71416 halobacteri
10	524	21.9	451	1 Q06022	Q06022 halobacteri
11	523.5	21.9	544	1 P71409	P71409 halobacteri
12	523.5	21.9	643	17 Q9HP05	Q9HP05 halobacteri
13	521.5	21.8	628	17 Q9HRN6	Q9HRN6 halobacteri
14	509.5	21.3	789	17 Q9HRA1	Q9HRA1 halobacteri
15	507.5	21.2	788	1 Q93643	Q93643 halobacteri
16	506	21.1	420	17 Q9HPW6	Q9HPW6 halobacteri

17	505	21.1	642	1 Q59634	Q59634 halobacteri
18	502.5	21.0	773	1 Q9UXS0	Q9UXS0 halobacula
19	499.5	20.9	627	17 Q9HS86	Q9HS86 halobacteri
20	497.5	20.8	419	1 Q93644	Q93644 halobacteri
21	497	20.8	536	17 Q9HPA2	Q9HPA2 halobacteri
22	475.5	19.9	537	16 Q9AB06	Q9AB06 caulobacter
23	452.5	18.9	423	17 Q9H0T5	Q9H0T5 halobacteri
24	439.5	18.4	423	1 Q93642	Q93642 halobacteri
25	419.5	17.5	559	16 Q9K659	Q9K659 bacillus ha
26	415	17.3	535	16 Q9A5Y0	Q9A5Y0 caulobacter
27	413	17.3	439	16 Q9KRF4	Q9KRF4 bacillus ha
28	409.5	17.1	632	16 Q9HW93	Q9HW93 pseudomonas
29	408	17.0	499	16 Q8UGL0	Q8UGL0 agrobacter
30	405.5	16.9	579	16 Q9K632	Q9K632 bacillus ha
31	405.5	16.9	632	2 Q32440	Q32440 pseudomonas
32	401	16.8	530	16 Q9X0W7	Q9X0W7 thermotoga
33	401	16.8	629	16 Q9HW91	Q9HW91 pseudomonas
34	400.5	16.7	553	16 Q9KNF6	Q9KNF6 vibrio chol
35	400	16.7	664	16 Q8RNV6	Q8RNV6 thermoaer
36	399.5	16.7	535	16 Q910R3	Q910R3 pseudomonas
37	399	16.7	702	16 Q97D01	Q97D01 clostridium
38	397	16.6	566	16 Q9X1E2	Q9X1E2 thermotoga
39	396.5	16.6	541	16 Q9HUP8	Q9HUP8 pseudomonas
40	396.5	16.6	834	17 Q29217	Q29217 archaeoglob
41	394	16.5	629	2 Q32443	Q32443 pseudomonas
42	393	16.4	667	16 Q9K617	Q9K617 bacillus ha
43	392.5	16.4	596	16 Q9KRJ7	Q9KRJ7 vibrio chol
44	391	16.3	667	2 Q929T0	Q929T0 bacillus ha
45	390	16.3	623	16 Q9K043	Q9K043 vibrio chol

ALIGNMENTS

RESULT 1	ID	P71412	PRELIMINARY:	PRT:	482 AA.
AC	P71412:				
DT	01-FEB-1997 (TREMUREL. 02, Created)				
DT	01-FEB-1997 (TREMUREL. 02, Last sequence update)				
DT	01-MAR-2002 (TREMUREL. 20, Last annotation update)				
DE	Transducer HCA protein.				
GN	HCA.				
OS	Halobacterium salinarum.				
OC	Archaea; Euryarchaeota; Halobacteriia; Halobacteriales;				
OC	Halobacteriaceae; Halobacterium.				
OX	NCBI_TaxID=2242;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	WEDLINE-96209786; PubMed-8643458;				
RA	Zhang W., Brooun A., McCandless J., Banda P., Alam M.;				
RT	"Signal transduction in the archaeon Halobacterium salinarum is				
RT	processed through three subfamilies of 13 soluble and membrane-bound				
RT	transducer proteins."				
RL	Proc. Natl. Acad. Sci. U.S.A. 93:4649-4654(1996).				
DR	EMBL; U75435; AAB17880.1; -.				
DR	HSSP; P02942; I007.				
DR	InterPro; IPR004089; Chntaxis_transd.				
DR	InterPro; IPR003660; HAMP.				
DR	InterPro; IPR000014; PAS_domain.				
DR	Pfam; PF000015; MCPsignal; 1.				
DR	SMART; SM00304; HAMP; 1.				
DR	SMART; SM00283; MA; 1.				
DR	SMART; SM00091; PAS; 1.				
SO	SEQUENCE 482 AA; 50872 MW; ECBEF79B3374CEC7 CRC64;				
QY	15 NGIDSHALADRIIGLDEAIAMRLSFTGIDDTMALAAEQPLFEATADAL-VTDFYDHLF 73				
	Query Match	25.5%	Score 609.5;	DB 1;	Length 482;
	Best Local Similarity	37.2%	Pred. No. 1.5e-18;		
	Matches 181; Conservative	70;	Mismatches 197;	Indels 39;	Gaps 12;

```

Db 19 NTLDPAF--VDADGAVVAM-----DDQIAL-----LETABEDAIQVTDIGERLN 63
Oy 74 SYERTODLANSRTKVEOLKETQAEYLLGIRGEYDETEYAQAORARIGKIHVDLGPVY 133
Db 64 D----DGSALAKKAVDTIDAHHY---DGYGLADESTALLTGVDYEDTYAGNTDLM 116
Oy 134 LGAYTRYTGLLDLADLDVADVRGEEAAAVDELVARFLPMLKL-----TFDQOI--- 184
Db 117 FIATPYHNGEERGVY-EIVQDR-SSSARYQSELQALFGLVDTLDAYDAGRDATAVDIA 174
Oy 185 AMDYIDSAQRLHDEIDSRQELANAVATHVEPLSSLEATSGDVAERTDTMAKRTDDV 244
Db 175 AEDTLDDDEYIQIGRLNTEFGDTLAAHTEVHNDAVRLERLAAQVSSAEIDELSTAAQ 234
Oy 245 DRMAVDSREISSVASEVEASTADVDRTSEDAEALAOGEAAADALATMTDIDEATD 304
Db 235 TIVSYVATEVEITSAVQELASTADEVDTSAERLADGSAASDAADMAADVATAAD 294
Oy 305 GVTAGVEOLGERAADVESYTGVIDDIAEQTMLALNASTIEAARAGEGFAVVADEVKA 364
Db 295 SVTSDEVALONRIEDIDEVVDYVTGIAEQTNMLALNASTIEAAPGEGEGFAVVADEVKA 354
Oy 365 LAESREOSTREVELVEQMAETEETVDOLDEVNORIGEVERVEEAMETLQETITAVE- 423
Db 355 LAEDAQSNAGHIESLSEIORDTADVTDLVYTTDRIEDAVAOVEDAMASFEEITVAVHA 414
Oy 424 --DAASGMOEVSTATDEQAVSTEEVAVDVGDVRAGEITAAALDDIDADATDOQVTEVEY 481
Db 415 TAERASNM--VSDATNEQAAAEELAAVDETADLADLITTAADVIVSOTEAQAMLHDL 472
Oy 482 RETVGRK 488
Db 473 DESVSEL 479

```

RESULT 2

Q9HP10 PRELIMINARY; PRT; 633 AA.

```

ID Q9HP10
AC Q9HP10;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Htr3 transducer.
GN HTR3 OR VNG1856G.
OS Halobacterium sp. (strain NRC-1).
OC Archaea: Euryarchaeota: Halobacteriia: Halobacteriales:
OC Halobacteriaceae: Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahalax G.G., Bergquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Shroogha J.,
RA Swartzell S., Welt D., Hall J., Dahl T.A., Welt R., Goo Y.A.,
RA Leithausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbacher T.A., Peck R.F., Pohlschroder M., Spudis J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarina S.;
RT "Genome sequence of Halobacterium species NRC-1."
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AJ245950; CAB82572.1; -.
DR HSSP; P02942; 10U7.
DR InterPro; IPR004089; Chmtaxis_transd.
DR InterPro; IPR003660; HAMP.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF00015; MCPsignal; 1.
DR PRINTS; PR00260; CHEMTRNSDUCR.
DR SMART; SM00304; HAMP; 2.
DR SMART; SM00283; MA; 1.
KW Complete proteome.
SQ SEQUENCE 633 AA; 66423 MW; C4D8022B66ECD0FF CRC64;

```

Query Match 24.9%; Score 597; DB 17; Length 633;
 Best Local Similarity 33.8%; Pred. No. 7, 1e-18;
 Matches 150; Conservative 76; Mismatches 158; Indels 60; Gaps 7;

```

Oy 42 IDDDTMAALAEOPLEFATADALVTFYHLESEYRT----QDLFANSRTKVEOLKETQA 97
Db 213 LDEEDLGAAGASLSQMHTRLELITDLDEAREDAEOTRKDAEARAASRLNERLRRA 272
Oy 98 EYLLGIRGEYDETEYAQAORARIGKIHVDLGLGPVYLGAYTRYTGLLDLADDDVADRG 157
Db 273 -----EYSDMAAAA-----GDLTR-----RLDEDVSEPM 299
Oy 158 EEAATAVDELVARFLPMLKLTFDQOIAMDYIDSYAQRRLHDEIDSRQELANAVATHVEA 217
Db 300 QDLAEAFNMMG-----DVEATL-----AQVRSIADAVD-----AASDVST 336
Oy 218 PLUSLEATSDQVAERTDTMAKRTDDQVDRMAVDSREISSVASEVEASTADVDRTSED 277
Db 337 SAAEIRASADQVSESVQDISADADQDRDLGTGDEVSTLSATVEIEIAASADVAETVNO 396
Oy 278 AEAALAOGEAAADLALATMTDIDEATDGVYTAGVEOLGERAADVESYTGVIDDIAEQTML 337
Db 397 AATESRGGELGEDAVAELELIEATDSAVERTYALEAVDAIGDVGITDIAEQTML 456
Oy 338 ALNASTIEAARAGEGFAVVADEVKALAEESREOSTREVELVEQMAETEETVDOLDEV 397
Db 457 ALNANIEAARADKSGGFAVVADEVKALDADEVKESATLETITLVDDQVADVADVADMSLE 516
Oy 398 NORIGEVERVEEAMETLQETITDAVEDAAAGMOEVSTATDEQAVSTEEVAVDVGDVRA 457
Db 517 GDRVDAGSEETIEAALALDDIGQVFAANGSVQISIDATDEQAASVEEVMTIDEVTDL 576
Oy 458 GEIAAALDDIADATDOQVTEVEY 481
Db 577 DRTATESQGVSAAEQAASVSEV 600

```

RESULT 3

Q9P9J0 PRELIMINARY; PRT; 805 AA.

```

ID Q9P9J0
AC Q9P9J0;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Chemotaxis transducer protein Bast.
GN BAST.
OS Halobacterium salinarum.
OC Archaea: Euryarchaeota: Halobacteriia: Halobacteriales:
OC Halobacteriaceae: Halobacterium.
OX NCBI_TaxID=2242;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=59;
RA MEDLINE=20138366; PubMed=10672186;
RA Kokoeva M.V., Oesterhelt D.;
RT "Bast, a membrane-bound transducer protein for amino acid detection in
RL Halobacterium salinarum."
RL Mol. Microbiol. 35:647-656(2000).
DR EMBL; AJ245950; CAB82572.1; -.
DR HSSP; P02942; 10U7.
DR InterPro; IPR004089; Chmtaxis_transd.
DR InterPro; IPR003660; HAMP.
DR InterPro; IPR004090; Me_chemotaxis.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF00015; MCPsignal; 1.
DR PRINTS; PR00260; CHEMTRNSDUCR.
DR SMART; SM00304; HAMP; 2.
DR SMART; SM00283; MA; 1.
SQ SEQUENCE 805 AA; 84831 MW; 089CA734D3F9DE3D CRC64;

```

Query Match 24.9%; Score 597; DB 1; Length 805;
 Best Local Similarity 33.8%; Pred. No. 9, 5e-18;

Matches 150; Conservative 76; Mismatches 158; Indels 60; Gaps 7;

QY 42 IDDDTMAALAEQPLFENTADATVDFYDHESTERT-----QDLEFANSTKVEQLKEFQA 97
Dd 385 LDEBLPGFAGSLSQMHTRLALITDDEAREDAEQTKDAEEARASERLNERLERRAA 444
QY 98 EYLLGLGCEYDTEYAAQARIRIGKIHVLGIPVYLGAYRYRTGLDLADLVADRG 157
Dd 445 -----EYSDMEAAAA-----GDLTR-----RLDEVDSEPM 471
QY 158 EEAANADELVARFLPMKLLTFDQOAMDTYIDSYAORLHDEIDROELANAVATHEA 217
Dd 472 QDIEAENDMWG-----DVEATL-----AQVRSIDAVD-----AASTDVST 508
QY 218 PLSLAETSDOVARERTDTRMARTDDQVDRMADVSREISSVASVEASTADVARTSED 277
Dd 509 SAAEIRSDQVSESVDISADADQQRDRIGTVGDEVYSLATVEEIAASADVAEYVNO 568
QY 278 AEALAOGEAADDALATMTDIDEATDGTAGVEQLGERADVESVTGVIDDIAEQTNL 337
Dd 569 AATSESGOELGEDAVALELRIEATPDSAVERTALAEVAIGVITVDIAEQTNL 628
QY 338 ALNASTAARAGEGEGFAVAVADEVKALAEESREOSTREVELVEQMAETEETVDQDDEV 397
Dd 629 ALNASTAARAGEGEGFAVAVADEVKALAEESREOSTREVELVEQMAETEETVDQDDEV 688
QY 398 NORIGEVEREEMETLOEITDAVEDAASGMOESTRTDQOAVSTEEVAMVQVDRA 457
Dd 689 GDRVAGSETTEALALADDDIGDOVEAANGSVOSTISDATTQOASTEEVMTIDEVTL 748
QY 458 GEIAAALDDIADATDOQVRYEEV 481
Dd 749 DRTATESQVSAABEQAAVSSEV 772

RESULT 4
Q9HR88 PRELIMINARY; PRT: 790 AA.
AC 09HR88;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DE Hcr18 transducer.
GN HTR18 OR VNC0812G.
OS Halobacterium sp. (strain NRC-1).
OC Archaea; Euryarchaeota; Halobacteriia; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RX SEQUENCE FROM N.A.
MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahaitas G.G., Bergquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sirogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithausen B., Keller K., Cruz R., Danon M.J., Hough D.W.,
RA Maddocks D.G., Jablonksi P.E., Krebs M.P., Angelone C.M., Dale H.,
RA Isenberger T.A., Peck R.F., Pohlschoder M., Spudich J.L., Jung K.-H.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RA Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
RL "Genome sequence of Halobacterium species NRC-1".
DR EMBL: AE005023; AAC19270.1; -.
DR HSSP: P02942; 10U7.
DR InterPro: IPR004089; CtmTaxis_transd.
DR InterPro: IPR003660; HAM.
DR InterPro: IPR004090; Me_Chemotaxis.
DR Pfam: PF00672; HAM; 1.
DR Pfam: PF00015; MCPsignal; 1.
DR PRINTS: PR00260; CHEMTNNSDCR.
DR SMART: SM00304; HAM; 2.
DR SMART: SM00283; MA; 1.
KW Complete proteome.
SQ SEQUENCE 790 AA; 84320 MW; 6227D876B5A35B85 CRC64;

Query Match 24.2%; Score 580; DB 17; Length 790;
Best Local Similarity 31.4%; Pred. No. 4,8e-17;
Matches 160; Conservative 84; Mismatches 190; Indels 76; Gaps 11;

QY 1 MSNDTLYTADVNRGIDGHALADRICGLDEAIEIRWLSFTGIDD--DTMALA--AEOP 55
Dd 286 MSDRREALAAGDIDTAIEETTRIDEVG-----ILRSFRIOELQTVAGQADALAEOD 339
QY 56 LFEATA-----DALVTDFD--HLESYRTQDLFANSTKYVQLKEQAEYLLG 102
Dd 340 -FDADALDKSVPGRGESLETHMHDLETAIDLDAQETAQSKREAGQSR-EAEALA- 396
QY 103 LGRGEYDTEYAAQARIRIGKIHVLGIPVYLGAYTRYTGLDLADLVADRGEEAA 162
Dd 397 -----AALESQAODIREYEHAD-----GDLQRLTDTDHESMA 432
QY 163 AVDELVARFLPMKLLTFDQOIAMDTYIDSYAORLHDEIDROELANAVA--THVEAPL 219
Dd 433 AL-----ATALNSLLELEGTIHRIGRSKVAESSDHITTS 470
QY 220 SLEATSDOVARERTDTRMARTDDQVDRMADVSREISSVASVEASTADVARTSED 279
Dd 471 EEVVRASQVSESVQEMKADARQONGIVQDVSDLETTDLATIEETIASSSDEVAAKSNDV 530
QY 280 ALAOGGEAADDALATMTDIDEATDGTAGVEQLGERADVESVTGVIDDIAEQTNL 339
Dd 531 SVGSGARARSODALEEMNAVDEQAKRTIAEMALDDEMTIEGEITTLIDDAEOTSMAL 590
QY 340 NASIEAARAGEGEGFAVAVADEVKALAEESREOSTREVELVEQMAETEETVDQDDEV 399
Dd 591 NASIEAARAGEGEGFAVAVADEIKSLSKETTEA--QEISLADQDSTDAVTDMQEGD 650
QY 400 RIGSEVEREEMETLOEITDAVEDAASGMOESTRTDQOAVSTEEVAMVQVDRA 459
Dd 651 RLSEKSTVTDTVETIDTIVRIEIRANGVQTI--TATDEQATTEEVYTNVDEVSIDD 710
QY 460 IAAALDDIADATDOQVRYEEVRYEYKLS 489
Dd 711 TTAAREDAANAABEQTASLTETVNRIDPLS 740

RESULT 5
O93775 PRELIMINARY; PRT: 452 AA.
AC O93775;
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DE Car protein.
GN CAR.
OS Halobacterium salinarum.
OC Archaea; Euryarchaeota; Halobacteriia; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=2242;
RN [1]
RX SEQUENCE FROM N.A.
MEDLINE=99164082; PubMed=10064582;
RA Storch K.F., Rudolph J., Oesterhelt D.;
RA "Car: A cytoplasmic sensor responsible for arginine chemotaxis in the
RA archaeon Halobacterium salinarum.";
RL EMBL J. 18:1146-1158(1999).
DR HSSP: P02942; 10U7.
DR InterPro: IPR004089; CtmTaxis_transd.
DR InterPro: IPR004090; Me_Chemotaxis.
DR Pfam: PF00015; MCPsignal; 1.
DR PRINTS: PR00260; CHEMTNNSDCR.
DR SMART: SM00283; MA; 1.
DR SMART: SM00091; PAS; 1.
DR TIGRfams; TIGR00229; sensory_box; 1.

SEQUENCE 452 AA: 49098 MM; 9244D900694681A CRC64;
Query Match 23.4%; Score 560.5; DB 1; Length 452;
Best Local Similarity 31.3%; Pred No. 1,6e-16;
Matches 159; Conservative 88; Mismatches 170; Indels 91; Gaps 13;
QY 10 TADVRNGIDGHALADRI-----GLDEAEIA-----WRISFTGIDDDTMAALAA 52
DB 5 SSDMGGEATGEHLADELCAYLGDNEDDGDDELQRLSRERDFWKHMFN-----QLVA 56
QY 53 EDP--LEFATADALVTD-----FYDHEST-----ERTDULFANSKYTYEQLEKTOAETL 100
DB 57 EYPEGGLTAAAGCTVTYHWNERSFDHKKMARSDALGEDASDVS---TAEE--SETLPEAV 111
QY 101 LGLGRGEYDTEYAARARIGKIHVYLGIPDYLGATRYTYGLDALADADVADRGEBA 160
DB 112 VRTGDTVEEED-----PHDVTDSLQYHGVPLRAPTG-----DVGSGT--V 152
QY 161 AAADVLARFLDMLKLTFFDQOIAMDYIDSYAQRHDEIDSRQELANAVATHEAPUS 220
DB 153 VDISKRYKN-----QRELHD-----LHETVSSNVEHLIS 182
QY 221 SLEATSDQVAERFDIMRARTDDQVRMADVSREISSVSASVEEVASTADVARTSFDAA 280
DB 183 ELSESIDEVGSRAETEAREAGEIEEMEGVADEVSNQSRATIEIASAEVSAQSRAD 242
QY 281 LAQGEFAAADA LATWTDIDEATDGTAGVTEQLGERAAVDESATGVIDIAEQTNMLALN 340
DB 243 RATEGQGTETALIDRMGAQGESAEKRNQDIDLTGQADEMSLIDAINIAQTNMLALN 302
QY 341 ASIEARAGEAGEGFAVVADEVKALAEESREOSTRYEELVEQMATEETVVDQLEVNOR 400
DB 303 ASIEARAGEAGEGFAVVADEVKALAEESREOSTRYEELVEQMATEETVVDQLEVNOR 362
QY 401 IEGEVRVEAEETLOETDAVEDAASGQEVSTATDQAVSTEEVAEAVDQVDRAGET 460
DB 363 IEGEVRVEAEETLOETDAVEDAASGQEVSTATDQAVSTEEVAEAVDQVDRAGET 422
QY 461 AAALDDIADATDQVATVEEVEETVQKL 488
DB 423 EDRLDLSQIASQSDHRAVEDMDDEL 450

RESULT 6
P71415 PRELIMINARY; PRT: 777 AA.
AC P71415: 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
GN Transducer Htd protein.
OS Halobacterium salinarum.
OC Archaea; Euryarchaeota; Halobacteriia; Halobacteriales;
CC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=2242;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FLX15;
RX MEDLINE=96209786; PubMed=8643458;
RA Zhang W., Brooun A., McCandless J., Banda P., Alam M.;
RT "Signal transduction in the archaeon Halobacterium salinarum is
processed through three subfamilies of 13 soluble and membrane-bound
transducer proteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:4649-4654(1996).
DR HSSP; P02942; 1007.
DR InterPro; IPR004089; Chttaxis_transd.
DR Pfam; PF00672; HAMF. 1.
DR PRINTS; PR01608; BACINVASINC.
DR SMART; SM00304; HAMF. 2.
DR SMART; SM00283; MA. 1.
DR SMART; SM00283; MA. 1.

SEQUENCE 777 AA: 82097 MM; 0B20AA6A593AA816 CRC64;
Query Match 22.6%; Score 540.5; DB 1; Length 777;
Best Local Similarity 32.4%; Pred No. 2,2e-15;
Matches 141; Conservative 74; Mismatches 171; Indels 49; Gaps 6;
QY 88 TVEQLKETQAEYLLGLGRGEYDTEYAAGRA-RIGKIHVYLGIPDYLGATRYTYGGLD 146
DB 320 TVRALNDLEAK-APALERGEYDIDLVAHVDELGRLEFAFSLRVAQVRAQSDANEQVD 378
QY 147 ALA-----DQVADRGEAAAVDELARFLP 173
DB 379 AEAARSEAAQAEPBAQAQAEAREESDAGAPPCDDRGVLDGADACVAAQDLTVR--- 435
QY 174 MLKLTFFDQIA-----MDYIDSYAQRHDEIDSRQELANAVA---THEAPLSLEATS 226
DB 436 ----LDADVEQAAMDNRATFEMADNATIAERKAFDEVAATSTDSNAAVEQIG 491
QY 227 QDVAERTDMRARTDDQVRMADVSREISSVSASVEEVASTADVARTSFEAALAQGE 286
DB 492 RQVSTAVGRIRDRADQDRDLAVASSETDEMSATIEEVARVAGVAETSORAAALGDDQ 551
QY 287 AAADALATMTDIDEATDGTAGVTEQLGERAADVESATGVIDIAEQTNMLALNASTEAA 346
DB 552 AAQAQAVVAQLEIEDETQAAATAVVDLEAKMSEITETVAITDIEQTNMLALNASTEAA 611
QY 347 RAGEAGEFAVVADEVKALAEESREOSTRYEELVEQMATEETVVDQLEVNORIGEGVE 406
DB 612 RADQGDGFAVVADEVKALAEESREOSTRYEELVEQMATEETVVDQLEVNORIGEGVE 671
QY 407 RVEAMETLOETDAVEDAASGQEVSTATDQAVSTEEVAEAVDQVDRAGET 466
DB 672 TVSETERSLSDIAGRIAEPTDQVETISNMDQAAVSVDVTAQVDAALGSETATEAES 731
QY 467 IADATDQVATVEE 481
DB 732 TRDAAAEQATTLSDV 746

RESULT 7
P71414 PRELIMINARY; PRT: 792 AA.
AC P71414: 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
GN Transducer Htc protein.
OS Halobacterium salinarum.
OC Archaea; Euryarchaeota; Halobacteriia; Halobacteriales;
CC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=2242;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FLX15;
RX MEDLINE=96209786; PubMed=8643458;
RA Zhang W., Brooun A., McCandless J., Banda P., Alam M.;
RT "Signal transduction in the archaeon Halobacterium salinarum is
processed through three subfamilies of 13 soluble and membrane-bound
transducer proteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:4649-4654(1996).
DR HSSP; P02942; 1007.
DR InterPro; IPR004089; Chttaxis_transd.
DR Pfam; PF00672; HAMF. 1.
DR PRINTS; PR01608; BACINVASINC.
DR SMART; SM00304; HAMF. 2.
DR SMART; SM00283; MA. 1.
DR SMART; SM00283; MA. 1.
DR SMART; SM00283; MA. 1.

Query Match 22.5%; Score 539; DB 1; Length 792;

Best Local Similarity 32.2%; Pred. No. 2.6e-15;
Matches 143; Conservative 76; Mismatches 163; Indels 62; Gaps 7;

```
OY 42 IDDDTMALAAEQPLEFETADALYDFDHESEYRT-----QDLFANSTKVEQLKETQA 97
D 374 LDEDPGAFGASLSQMH-RLERLITDDEARDAEQTRKDAEENAAERLNERLERRAA 432
OY 98 EYLLGLGGEVDTEYAAORARIGKIHVILGLCPDVYLCAATRYRYTGLDALADADVADRG 157
D 433 -----EYSDMAAAA-----GDLR-----RLDEVDSDGM 459
OY 158 EAAAAVDELVARFLPMLKLLTFDQIAMDYIYSYAQRHLDEIDSRQELANNAVATHEA 217
D 460 QDIAEAFNDMKGDV-----EATLAQVRISADAVDASTVDHDAEIR-----501
OY 218 PLSSLEATSDVAERTDTMRARTDQVVRMADVSREISSVSASVEVASTADVARTSED 277
D 502 -----SASDDVSESVQDISADAEQRDRLGTGVGEVHALSTVBDIARPTTSPSTVNO 555
OY 278 AEALAQGEAAADALATMTDIDEATDGTAGVEQLGERAADVESVTGVIDDIABOTNML 337
D 556 AATSEERQELGEDAVALEERTEATADSAVERVTALEAVDAIGHVYITDIAEQTNML 615
OY 338 ALMASIARAGEAGEGFAVVADEVKALAEESREOSTRVEELVEQMAETEETVDQLDEV 397
D 616 ALNANIEAARADKSGDGFVAVVADEVKDLADEVKESATELETLLVDQVADVADTVADESEL 675
OY 398 NORICEGVERVEAMETLOETIDAVEADAASGQOEVTATDEQAVSTEEVAEKVGDVDDRA 457
D 676 GDRVAGSETTEALAAVDDIGDOVEAANGSVOSISDATDEQAASTEVEVTMIDEVTDLS 735
OY 458 GEIQAALDDIADATDOQVRYVEEV 481
D 736 DRTATESQVSAAEQAASVSEV 759
```

RESULT 8

```
O9HOX8 PRELIMINARY: PRT: 636 AA.
AC O9HOX8.
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Htr15 transducer.
GN HTR15 OR VNC09586.
OS Halobacterium sp. (strain NRC-1).
OC Archaeae; Euryarchaeota; Halobacteriia; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahaliras G.G., Bergquist B., Pan M.,
RA Shukla H.D., Laskey S.R., Balliga N.S., Thorsson V., Sprooga J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenberger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1."
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AE005032; AAG19381.1;
DR HSSP; P02942; 1007.
DR InterPro; IPR004089; CtmTaxis_transd.
DR InterPro; IPR004090; Me_chemotaxis.
DR InterPro; IPR000700; PAS_assoc.C.
DR InterPro; IPR000014; PAS_domain.
DR Pfam; PF000015; MCPsignal; 1.
DR Pfam; PF00988; PAS; 1.
DR PRINTS; PR00260; CHEMTRNSDCR.
DR SMART; SM00283; MA; 1.
DR SMART; SM00091; PAS; 1.
```

DR TIGRFAMS; TIGR00229; sensory_box; 1.

KW Complete proteome.

SO SEQUENCE 636 AA; 67348 MW; 98D2E41F295D4D14 CRC64;

Query Match 22.4%; Score 535.5; DB 17; Length 636;

Best Local Similarity 29.7%; Pred. No. 2.8e-15;

Matches 149; Conservative 99; Mismatches 191; Indels 63; Gaps 11;

```
OY 17 IDGHALADRIGLDEAEIMRLSFGICIDDTMAALAAEQPLEFAT-----ADALYND-----67
D 165 VDDDLDSIGMP-----VFMLDATG-----SVAAWNTIEELTGCSKADAAQMGHASEA 214
OY 68 FYDHLSEYRTODLFFANSTKVEQLKETQAEYLGLGGEVDTEYAAORARIGKIHVILG 127
D 215 FYPPDRRKVKTLD-----KYIESPRSA-----EPDIESEKAQLYNTSYM- 258
OY 128 LGPDVYLCAATRYRYTGLDALAD-----VADRGEA--AAAVDELVARFLPML- 175
D 259 ----TDQGTDRHRIFFKASPIFDGDELLAAVEETEDRTEDVRRADAAVEELVDELSTYID 314
OY 176 -----KLTFDQIAMDYIYSYAQRHLDEIDSRQELANNAVTHVAPLSLEATSO 227
D 315 ALSSGOLSKRASFEHEGITNEQLVSVALNGMAQDFRLVGQVGGQTELDADTERRATA 374
OY 228 DVAERTDTMRARTDQVVRMADVSREISSVSASVEEVASTADVARTSEDAEALAQGEA 287
D 375 DATDIADTV-----SSQENMLSSAANEMENFSAHQEVAASSQVAAADQADAAESGLE 430
OY 288 AADALATMTDIDEATDGTAGVEQLGERAADVESVTGVIDDIABQTNMLANASIEAR 347
D 431 ASEGANQATNEVIDYISDDLMEVSXLESRMDEI EYVEVIAEVADQTNILANANIEAR 490
OY 348 AGEAGEGFAVVADEVKALAEESREOSTRVEELVEQMAETEETVQDLDEVNORICEGVER 407
D 491 AGEAGSGFAVVADEVKELANETREHETRIAGSI SDVQOQANETVAIVESHQHRRAGDE 550
OY 408 VEEAMETLOETIDAVEADAASGQOEVTATDEQAVSTEEVAEKVGDVDDRAGEIAAALDDI 467
D 551 IDDALITALBEIATSVDEAATGITEVARRANDEQASITVEDVYITIEDVQQAEEBAASARI 610
OY 468 ADATDOQVRYVEEVRETVKLS 489
D 611 VSATQEGSTAVASQLSERVDKLT 632
```

RESULT 9

```
P71416 PRELIMINARY: PRT: 804 AA.
ID P71416.
AC P71416.
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Transducer HCF protein.
GN HCF.
OS Halobacterium salinarum.
OC Archaeae; Euryarchaeota; Halobacteriia; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=2242;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=FLX15;
RX MEDLINE=96209786; PubMed=8643458;
RA Zhang W., Brooun A., McCandless J., Banda P., Alam M.;
RT "Signal transduction in the archaeon Halobacterium salinarum is
RT processed through three subfamilies of 13 soluble and membrane-bound
RT transducer proteins."
RL Proc. Natl. Acad. Sci. U.S.A. 93:4649-4654(1996).
DR EMBL; U75439; AAB17884.1;
DR InterPro; IPR004089; CtmTaxis_transd.
DR InterPro; IPR003660; HAMP.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF00015; MCPsignal; 1.
DR SMART; SM00304; HAMP; 2.
```

DR SMART; SM00283; MA; 1.
SQ SEQUENCE 804 AA; 84705 MW; 0D57284683BFD18B CRC64;

Query Match 22.1%; Score 529; DB 1; Length 804;
Best Local Similarity 30.6%; Pred. No. 6.9e-15;
Matches 139; Conservative 76; Mismatches 159; Indels 80; Gaps 9;

```

OY 89 VEOLKTEQAEVYLGLGRGEVDETEYAQR-----ARIGKHIDVYGLGPDVYLGATRYR 140
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 351 VQSLASAAAE-----LEAGNVDDVATSRDEIGQLFASIGSMRDALVTQAE---AAREDA 403
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 141 YVGLLDALD---DVAVDGEEPAAAVDELVARFLPMKLKLTFFDOQIAMDTYIDSYAQR 196
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 404 TEAQOQAEADARAEADARAEADAKADAEALAAE-----LEAQAEER 444
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 197 LHDEI-----DSR-----OELANAVAT---HY 215
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 445 YSDVIGGVADGDLTRRMPADDTDNEAMAAIAMSFNDSQMEHTIIDIQERADAVATASEEA 504
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 216 EAPLSLEATISQVARTDTMRARTDDQVDRMADVSRREISSVASYEEVASTADVRYRTS 275
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 505 EVGAADAERASGVSESVQEIAGADEQRNMLDTVSGEMTDLSAIEVYAS-ADSVAEHS 563
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 276 EDAAEALAOGEAAADALATMTDIDCATGCTAGVQGLGERADVSVTGVIDDIABEQTN 335
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 564 HQTAELARAGEQFAEDAEIERSLTVQEAIDATVQNVREALDOMAEISEIYDLISDIAEQTN 623
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 336 MLTANLSTAEARAGEGFAVVADEVKALAEESREOSTRVEELVQMOAETEETVDQDLD 395
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 624 MLTANLSTAEARAGEGFAVVADEVKALAEESREOSTRVEELVQMOAETEETVDQDLD 683
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 396 EVNQRIGEVEREAMETLQETITDAVEDAASGMOEVSTATDEQAVSTEEVAMVDGVD 455
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 684 AAESMDAGIDAEVVDFAFTAVSDSHSDDETQVGEISPTDQAASTEAVSMTEEVAD 743
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 456 RAGEIAAALDDIDATDQOVRTVEEVRVYGRKS 489
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 744 LSDSTAGEASVATAEQAASMSISDTVESLS 777
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 10

006022 PRELIMINARY; PRT; 451 AA.

```

AC 006022;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Soluble transducer protein Htl.
OS Halobacterium salinarum.
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=2242;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FLX15;
RX MEDLINE=97284501; PubMed=9139915;
RA Brooun A., Zhang W., Alam M.;
RT "Primary structure and functional analysis of the soluble transducer
   protein Htlxi in the Archaeon Halobacterium salinarum.";
RL J. Bacteriol. 179(12963-2968(1997)).
DR EMBL; U74668; AAC5264.1; -.
DR HSSP; P02942; 10U7.
DR InterPro; IPR004089; Chmtaxis_transd.
DR InterPro; IPR000014; PAS_domain.
DR Pfam; PF000015; MCPsignal; 1.
DR SMART; SM00283; MA; 1.
DR SMART; SM00091; PAS; 1.
DR TIGRFAMs; TIGR00229; sensory_box; 1.
SQ SEQUENCE 451 AA; 49122 MW; 6DA9E4FF8139DA87 CRC64;

```

Query Match 21.9%; Score 524; DB 1; Length 451;
Best Local Similarity 30.9%; Pred. No. 5.6e-15;
Matches 158; Conservative 87; Mismatches 168; Indels 98; Gaps 16;

```

OY 10 TADVRNGIDGHALADRI-----GLDEAEIA-----WRISFTGIDDDTMAALAA 52
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 5 SSDMGGEATGEHIADELCEAYLGDNEDGDELQRLSRERDMKMFN-----QLVA 56
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 53 EDP--LFEATADALVTD---FYDILEST-----ERTDDLPANSTKYVEQLKETQAEVL 100
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 57 EYPEGILITAAAGTAYTHMNERFSIHMKMARSDALGEDASDVFS---TAEE--SETLPEAV 111
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 101 LGIGREGEVTEYAQAARIGKIHVDVGLGAPDVYLGATRYVYVGLLDALADVDVADRGEEA 160
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 112 VRTGDTVEEEE-----PHDVPTDSLCOYHGVPLRAPPTG-----DVVGSFG--V 152
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 161 AAAYDELVARFLPMKLKLTFFDOQIAMDTYIDSYAQRHLEIDISROELANAVATHEAPLS 220
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 153 VPDISEKYN-----QRLHD-----LHFTVSINVEHLS 182
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 221 SLEATSDV---AERTDTMRARTDDQVDRMADVSRREISSVASYEEVASTADVRYRTS 277
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 183 ELSESIDKVGSEFAEETEPFAGK---EIERMEGFADKVSQSAIIEBIASAEVVSQASOR 239
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 278 AEALAOGEAAADALATMTDIDCATGCTAGVQGLGERADVSVTGVIDDIABEQTN 337
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 240 AQRATAGEGQTAEATDRMGAVQESAEVNDITGLTSQADESELIDAIINDIADQTNML 299
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 338 ALNASTAEARAGEGFAVVADEVKALAEESREOSTRVEELVQMOAETEETVDQDLD 397
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 300 ALNASTAEARAGEGFAVVADEVKALAEESREOSTRVEELVQMOAETEETVDQDLD 359
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 398 NQRIGEVEREAMETLQETITDAVEDAASGMOEVSTATDEQAVSTEEVAMVDGVD 457
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 360 TTEIEEATVAVRTTDSLOEIRNAVDETATGVAEVAGARD-HAASVEQVAATTTDEAVDKL 418
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 458 GETIAALDDIDATDQOVRTVEEVRVYGRKL 488
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 419 TELEDRLDNLQSASQHRVAIEIDMVDEL 449
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 11

P71409 PRELIMINARY; PRT; 544 AA.

```

AC P71409;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Transducer Htl protein.
GN Htl.
OS Halobacterium salinarum.
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=2242;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FLX15;
RX MEDLINE=96209786; PubMed=8643458;
RA Zhang W., Brooun A., McCandless J., Banda P., Alam M.;
RT "Signal transduction in the archaeon Halobacterium salinarum is
   processed through three subfamilies of 13 soluble and membrane-bound
   transducer proteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 93(4649-4654(1996)).
DR EMBL; U53365; AAB17519.1; -.
DR HSSP; P02942; 10U7.
DR InterPro; IPR004089; Chmtaxis_transd.
DR InterPro; IPR003660; HAMP.
DR Pfam; PF000015; MCPsignal; 1.
DR SMART; SM00304; HAMP; 1.
DR SMART; SM00283; MA; 1.
SQ SEQUENCE 544 AA; 56922 MW; 663DBD0EA43AFC71 CRC64;

```

Query Match 21.9%; Score 523.5; DB 1; Length 544;
Best Local Similarity 29.4%; Pred. No. 7.4e-15;
Matches 141; Conservative 93; Mismatches 195; Indels 51; Gaps 10;

28 IDEAEIANRLSFTGIDDDTMMALAAEOPLFEATADALVTFD-----YDHLESERYTODLFA 83
DB LEKAVAGTQI-FNLATGPLMATTVAVFALIGTATLAIYEDMALVEERQAPQAEHEERA 122
QY NSTKIVEOLKETQAEYLLGLGRGEYDT-EYAQAQRARICKIDVLCGLPDVYIGATRYYT 142
DB EBERAREKAEQKQAE-----AERQTAEOQSOTGRGTQREIEQLAAD--LESQATEYV 173
QY GLDLDLAD-----DIVADRGE--EAAAANDELARFLPMKLTLPFOQIAMDYIDYS 193
DB ATLEASOGDLTARVDATDTNNAIEAVATVNDM-----LTYMERTI----- 215
QY AORLDEIDSROELANAVATHEAPLSLEATSQDAERTDMRARTDDQVRMADVSRE 253
DB 216 -----DELQGFSTNTTASREATGAKETQVASFVSEVQIAGTDDQRLQLESVAEE 270
QY 254 ISSVASVEEVASTADVDVRRTSSEDAEALAOGEAAADALATMTDIDEATDGVAGVEQL 313
DB 271 MDSYATVEEVAATAQSVADTADTADTAGKQTAEDDAIDAIDAVQETMQTTVANVDAL 330
QY 314 GERADVESYVGVIDDIAEQTMMLNLSIEARAGEA-----GEGFAYVADVEYKALAES 369
DB 331 EDLTTEIDIAELISDIEQTMMLNLTENARAGSGGCTNGDFAVAVADEYKELATES 390
QY 370 REQSTRVEELVEMQAEETEVYDQLEEVNQRIGEGVEEEMETLQETTDVADEDAASGM 429
DB 391 QRSADINELIEVOSQRTATVEELIRVAEQRYNDGAAVEEYVDAGFATVEMIQETTDGV 450
QY 430 QEVSTATDEQAVSTEVAEMVDGVDRAGEIAALDDIADATDQVRYVEEYRETVGKLS 489
DB 451 QEISQAMDEQAQRSEYVSVVDIATISQATPADRAENVSAASEQGTASTITEVTSQSLA 510

RESULT 12

Q9HP05 PRELIMINARY: PRT: 643 AA.
AC Q9HP05;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Htr8 transducer.
GN HTR8 OR VNG1523G.
OS Halobacterium sp. (strain NRC-1).
OC Archaea: Euryarchaeota: Halobacteriia: Halobacteriales:
CC Halobacteriaceae: Halobacterium.
CX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Shrogha J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leitbauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Madlocks D.G., Jablonski P.E., Krebs M.P., Angvine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AE005065; AAC19812.1; -
DR HSSP; P02942; 10U7.
DR InterPro: IPR004089; Cmltaxis_transd.
DR InterPro: IPR003660; HAMP.
DR InterPro: IPR004090; Me_chemotaxis.
DR Pfam; PF00015; MCPsignal; 1.
DR PRINTS; PR00260; CHEMTRNSDCR.
DR SMART; SM00304; HAMP; 1.
DR SMART; SM00283; MA; 1.
KW Complete proteome.
SQ SEQUENCE 643 AA; 67290 MW; 1D53F2E18BA02481 CRC64;

Query Match 21.9%; Score 523.5; DB 17; Length 643;
Best Local Similarity 28.3%; Pred. No. 9e-15;

Matches 143; Conservative 92; Mismatches 200; Indels 71; Gaps 6;
QY 8 LVTAVRNGICSHALADDTGIDEAIANRLSTSEITDDDTMMALAAEQPLFEATADALVTD 67
DB 156 LVYVFTGCVFQMINAERYNHTTAIINPWVGSHGAPVLLAG-----ALMAN 205
QY 68 FYDHLSEYERTODLFANSTKTYEQLEKQAEYLLGLGRGEYTEVA----- 113
DB 206 WYSTERSRNASKRLREARQKQVDELEAR-----QAEIEAEKAEKRLKADAEARE 259
QY 114 ---AQRARICKIDVLCGLPDVYIGATRYYTGLDLDAL-DVADRGEAAAANDELVA 169
DB 260 AAEOQREVAALNLERLTANTYGAAMARAAGDLSVRDPVENDMAAIAASFENMU- 318
QY 170 RLPMKLTLPFOQIAMDYIDYSQRLHDEIDFQELANAVATV-----EAPLSSL 222
DB 319 -----DETETFTIREIOAASDVAAASEDADGAYEI 349
QY 223 EATSDVAERTDMRARTDDQVRMADVSREISVSASVEEVASTADVRRTSSEDAEALA 282
DB 350 EDASQVSETVQEIAGADEQREKLETYSGETPDLASAIIEVPAASADSVAERSHETAAVA 409
QY 283 QOGEAAADALATMTDIDEATDGVAGVEQLGERADVESYVGVIDDIAEQTMMLNLS 342
DB 410 GDEQTAQBAIDASRTVQSAVESVQVNEALDDQLAEISEIYDLISDVAEQTMMLNLAN 469
QY 343 IEAARAGEGEGFAYVADVEYKALAESREQSTRVEELVEMQAEETEVYDQLEEVNQRIG 402
DB 470 IEAARADSGDGFAYVADVEYKALAEETRASADIEALVAADIDQAGVATVTEARTDESQV 529
QY 403 EGVEEYEMETLQETTDVADEDAASGMQEVSTATDEQAVSTEVAEMVDGVDRAGEIAA 462
DB 530 DAISAVDAVVDAGFVGAENAEETDGTGVDEISTITTDQAASTEAVSMIAEVSIDISTATRA 589
QY 463 ALDDIADATDQVRYVEEYRETVGKL 488
DB 590 DAQASTAAEQTTAATISENTAAL 615

RESULT 13

Q9HRN6 PRELIMINARY: PRT: 628 AA.
AC Q9HRN6;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Htr16 transducer.
GN HTR16 OR VNG0614G.
OS Halobacterium sp. (strain NRC-1).
OC Archaea: Euryarchaeota: Halobacteriia: Halobacteriales:
CC Halobacteriaceae: Halobacterium.
CX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Shrogha J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leitbauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Madlocks D.G., Jablonski P.E., Krebs M.P., Angvine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AE005010; AAC19122.1; -
DR InterPro: IPR004089; Cmltaxis_transd.
DR InterPro: IPR003660; HAMP.
DR InterPro: IPR004090; Me_chemotaxis.
DR Pfam; PF00672; HAMP; 1.
DR PRINTS; PR00260; CHEMTRNSDCR.
DR SMART; SM00304; HAMP; 2.

[illegible]

Query Match	21.3%	Score 509.5	DB 17	Length 789
Best Local Similarity	30.6%	Pred. No. 4.5e-14		
Matches 144	Conservative 80	Mismatches 165	Indels 81	Gaps 8
DR Pfam: PF00015; MCPsignal; 1.				
DR PRINTS: PR00260; CHEMTRNSDUCR.				
DR SMART: SM00304; HAM; 2.				
DR SMART: SM00283; MA; 1.				
KW Complete Proteome.				
SO SEQUENCE 789 AA; 83926 MW; ABE57215065BF4B5 CRC64;				
QY 21 ALADRIGLDEAEIARLSTFTGIDDDTMAALAEOPLFETADALVTDFYHLESYERTOD 80	1	1	1	1
DB 361 SIQDRI--GEAF-----AAYEANAARAEAEELTADNEDQAQAEERAK- 401				
QY 81 LFANSTKTVIEQLKETQAEYELGLGGRGYDTEYAAQARIGKIHVYLGCPDYVLCAYTRY 140	1	1	1	1
DB 402 --ATAEASERLQERRAD----- 417				
QY 141 YTGLLDALADDDVADRGEEAA--AAVDLVARFLPMLKILTFDDQIAMDITYIDSYAQRH 198	1	1	1	1
DB 418 YSEVQAIVADGDLTERLEDDADEEMRAVATEFNML-----DGLAETIAQVAGFADEVA 472				
QY 159 DEIDSRQELANVATFHEVAPLSLSLATSODVAERTDTMFARTRDDQVDRADVSREISSVS 258	1	1	1	1
DB 473 DE-----TLQVATGAE----EIEITSQVYSERIQELINDGAIQGHDDLERRAGEMDEL 521				
QY 259 ASVEEVASTADDPRTSEDAEALAAQGEAAMADALATMTDIDEATDGVYTAGVEQLGERAA 318	1	1	1	1
DB 522 ASIQEVAASSTAVAEFTADADAVERGEGRDAESALIDMAEISLSADAVDQILALQERKS 581				
QY 319 DVEASTGVYDIDIAEQTNMLALNASTIEARAGAGSGFAVADVYALAEESNEOSTRYE 378	1	1	1	1
DB 582 DGDIDIEFTTDAEQTNMLALNASTIEARAGAGSGFAVADVYALAEESNEOSTRYE 641				
QY 379 LVEQQAETEEVDDQLDEVNORIGGVEEVEEAMETLOETIPAVDPAASGMOEVSTATDE 438	1	1	1	1
DB 642 EIQAVQAEITDELVDIRATSEHIDGVSIVYERAAAAIEPTVAIEDANHGIOEIDATED 701				
QY 439 QAVSTEVAEMVGVDDRAGEITMAALDDIADATDOQVTRVEEVRTYGL 488	1	1	1	1
DB 702 QADATQSVRYRVDVADISQHTDEDAEQVSAAEQGSASVAIIANSADL 751				
RESULT 15				
093643 PRELIMINARY; PRT; 788 AA.				
AC 093643				
DT 01-MAY-1999 (TREMBLrel. 10, Created)				
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)				
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)				
DE Transducer HtrVI.				
GN HTRVI.				
OS Halobacterium salinarum.				
OC Archaea: Euryarchaeota; Halobacteria; Halobacteriales;				
OC Halobacteriaceae; Halobacterium.				
OX NCBI_TaxID=2242;				
RN [1]				
RP SEQUENCE FROM N.A.				
RA BROUDEN A., Villablanca F., Freltas T., Hou S., Alam M.,				
RT "Structure, Function and Evolution of Transducer Family from the				
RT Archaeon Halobacterium salinarum.";				
RL J. Bacteriol. 0:0-0(1997).				
DR EMBL: AF036231; AAD02052.1; "				
DR HSSP: P02942; 10U7				
DR InterPro: IPR004089; Clmntaxis_transd.				
DR InterPro: IPR00360; HAM.				
DR InterPro: IPR00490; Me_chemotaxis.				
DR Pfam: PF00672; HAM; 1.				
DR Pfam: PF00015; MCPsignal; 1.				
DR PRINTS: PR00260; CHEMTRNSDUCR.				
DR SMART: SM00304; HAM; 2.				
DR SMART: SM00283; MA; 1.				

SO SEQUENCE 788 AA; 83868 MM; C0D0839DDDFC6CE CRC64;
Query Match 21.2%; Score 507.5; DB 1; Length 788;
Best Local Similarity 30.9%; Pred. No. 5.4e-14;
Matches 143; Conservative 79; Mismatches 160; Indels 81; Gaps 8;
QY 21 ALADRGDEAEIARLSTFGIDDDTMAALAEQPLFEATADALVTDFYDHLSEYERTOD 80
Db 361 SLDRI--QEE-----AAYEANAAKAEAEELRTDAEDQAEEAKK- 401
QY 81 LFANSTKTVQLKETOAEYLLGLGGEYDTEYAAQRARIGKIHVYLGIPDVLGAYTRY 140
Db 402 --ATAEASERLQERAD----- 417
QY 141 YTGLLDALADVDVADRGEAA--AAVDLVARFLPMLKLLTFDQIAMDYIYSTAQRLL 198
Db 418 YSEWQAQAVADDDLTERLDEADDEEMRAVATEFNML-----DGLEATIAQVAGFADEV 472
QY 199 DEISRQELANAVATHVEAPLSLEATSQDVAERTDTMRARTDDQVDRMADVSRREISSVS 258
Db 473 DE-----TIQVATGAE---EIEFTSQVSERIQETADGAIQOHDDLERRAGEMDEL 521
QY 259 ASVEEVASTADDVRRTSSEDAEALAOQGEAADDALATMTDIDEATDGYTAGVEQLGERAA 318
Db 522 ASIQEVAASATVAETADAVERGEGRDAAESATIDMAEIESLSADAVDQILALQERMS 581
QY 319 DVESYTGVIDIAEGTNMLNASTIARAGEAGGFAVVADEVKALAESEHREOSTRYEE 378
Db 582 DIGIIEFITDIAEQTNMLNANIEARADKDGGFVAVANEVYDLAEETKQAADIES 641
QY 379 LVEQMAETEETVDQLDEVNORIGEGEVEEAMETLOEITDAVEDAASGMQEVSTATDE 438
Db 642 EIQAVQAEETDETVDIRATSEHIDGVSTVERAAAATEDVDALEDANHGIOEISDATED 701
QY 439 QAVSTEYVAEMVDGVDRAGEIAAALDDIADATDQVRTVEEV 481
Db 702 QADATQSVRRVDVADISQHVTEDAEQVSAAEQASVAEI 744

Search completed: January 2, 2003, 12:42:42
Job time : 38 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2003, 12:36:46 ; Search time 14 Seconds

(without alignments)
1448.709 Million cell updates/sec

Title: US-09-455-978b-2

Perfect score: 2394
Sequence: 1 MSNDNDLTVTADVRNGIDGH.....ATDQVRYVEEVRVETGKLS 489

Scoring table: BLOSUM62
Gap 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2394	100.0	489	1	HMAT_HALSA
2	2360	98.6	489	1	HMAT_HALNI
3	643	26.9	481	1	HMAT_HALSA
4	643	26.9	481	1	HMAT_HALSA
5	578	24.1	778	1	HMAT_HALSA
6	573	23.9	778	1	HMAT_HALSA
7	555	23.2	810	1	HMAT_HALSA
8	554	23.1	810	1	HMAT_HALSA
9	537.5	22.5	534	1	HMAT_HALSA
10	521	21.8	545	1	HMAT_HALSA
11	492	20.6	433	1	HMAT_HALSA
12	464	19.4	763	1	HMAT_HALSA
13	463.5	19.4	763	1	HMAT_HALSA
14	461	19.3	764	1	HMAT_HALSA
15	449.5	18.8	535	1	HMAT_HALSA
16	412.5	17.2	432	1	HMAT_HALSA
17	388.5	16.2	662	1	HMAT_HALSA
18	386.5	16.1	662	1	HMAT_HALSA
19	379	15.8	662	1	HMAT_HALSA
20	359.5	15.0	547	1	HMAT_HALSA
21	359.5	15.0	1000	1	HMAT_HALSA
22	354.5	14.8	548	1	HMAT_HALSA
23	351	14.7	548	1	HMAT_HALSA
24	350.5	14.6	573	1	HMAT_HALSA
25	349	14.6	553	1	HMAT_HALSA
26	348.5	14.6	553	1	HMAT_HALSA
27	342.5	14.3	657	1	HMAT_HALSA
28	335.5	14.0	661	1	HMAT_HALSA
29	334	14.0	661	1	HMAT_HALSA
30	333.5	13.9	654	1	HMAT_HALSA
31	326.5	13.6	668	1	HMAT_HALSA
32	323.5	13.3	668	1	HMAT_HALSA
33	318.5	13.3	557	1	HMAT_HALSA

34	318	13.3	533	1	MCP2_RHIME
35	313.5	13.1	557	1	MCP2_CLOM
36	311.5	13.0	512	1	MCP2_ENTRAB
37	309	12.9	756	1	YAS1_RHISN
38	308	12.9	506	1	AER_ECOLI
39	304	12.7	533	1	MCP4_ECOLI
40	299.5	12.5	845	1	YAF4_RHISN
41	286	11.9	620	1	TCPI_VIBCH
42	271.5	11.3	417	1	FRZC_MYXXA
43	202.5	8.5	955	1	KINL_LEICH
44	196.5	8.2	1972	1	MYHB_HUMAN
45	195.5	8.2	1972	1	MYHB_RABIT

ALIGNMENTS

RESULT 1	ID	HMAT_HALSA	STANDARD	PRT	489 AA
AC	P71413				
DT	15-JUN-2002	(Rel. 41, Created)			
DT	15-JUN-2002	(Rel. 41, Last sequence update)			
DT	15-JUN-2002	(Rel. 41, Last annotation update)			
DE	Heme-based aerotactile transducer hemAT (Transducer HCB protein).				
GN	HEMAT OR HTB.				
OS	Halobacterium salinarum.				
OC	Archaea: Euryarchaeota: Halobacteriia: Halobacteriales:				
CC	Halobacteriaceae: Halobacterium.				
OX	NCBI_TaxID=2242;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=FLX15;				
RX	MEDLINE=96209786; PubMed=8643458;				
RA	Zhang W., Brooun A., McCandless J., Bahda P., Alam M.;				
RT	"Signal transduction in the archaeon Halobacterium salinarum is				
RT	processed through three subfamilies of 13 soluble and membrane-bound				
RT	transducer proteins.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 93:4649-4654(1996).				
RN	[2]				
RP	FUNCTION.				
RX	MEDLINE=20140131; PubMed=10676961;				
RA	Hou S., Larsen R.W., Boudko D., Riley C.W., Karatan E., Zimmer M.;				
RT	Oridal G.W., Alam M.;				
RT	"Myoglobin-like aerotaxis transducers in Archaea and Bacteria.";				
RL	Nature 403:540-544(2000).				
CC	-1- FUNCTION: Heme-containing signal transducer responsible for				
CC	aerotaxis, the migratory response toward or away from oxygen.				
CC	-1- SUBUNIT: Homotrimer (Probable).				
CC	-1- SIMILARITY: BELONGS TO THE CHEMOTAXIS SENSOR TRANSducers				
CC	FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
CC	EMBL: U75436; AAB1781.1; -				
DR	HSSP: P02942; 1007.				
DR	InterPro: IPR004089; Chmtaxis_transd.				
DR	Pfam: PF00015; MCPsignal; 1.				
DR	SMART: SM00283; MA; 1.				
KW	Transducer; Heme;				
SO	SEQUENCE 489 AA; 52816 MW; 84572636CF616BFF CRC64;				
Query Match	100.0%; Score 2394; DB 1; Length 489;				
Best Local Similarity	100.0%; Pred. No. 7e-97;				
Matches	489; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1. MSNDNDLTVTADVRNGIDGHALADRIIGLDEAEILAMRLSFTGIDDTMAALAEPLFEAT 60				

```

|||||
Db 1 MSNDNDTLVTADVRNGIDGHALADRIGLDEAEIAMRLSFTGIDDDPMALAAEQPLEFAT 60
QY 61 ADALVTDFDHLSEYERTQDLPFANSTKTYBQLEKTAQAEYLLGREGYDTEYAQARATG 120
Db 61 ADALVTDFDHLSEYERTQDLPFANSTKTYBQLEKTAQAEYLLGREGYDTEYAQARATG 120
QY 121 KIHVDVGLGPDVYLGAATRYTGTGLDALADDDVADGEEAAAANDELVARFLPMLKLTFF 180
Db 121 KIHVDVGLGPDVYLGAATRYTGTGLDALADDDVADGEEAAAANDELVARFLPMLKLTFF 180
QY 181 DQOIAMDVTYIDSYAQRLLHDEIDSRQELANAVATHVAPLSLEATSDQVAERTDTMARAT 240
Db 181 DQOIAMDVTYIDSYAQRLLHDEIDSRQELANAVATHVAPLSLEATSDQVAERTDTMARAT 240
QY 241 DQOVDNRADVSREISSVSASVEEVAATADVRRTSDEAEALAOOGGEAADALATMTDID 300
Db 241 DQOVDNRADVSREISSVSASVEEVAATADVRRTSDEAEALAOOGGEAADALATMTDID 300
QY 301 EATDGTAGVEQJGERAADVESYTGVIDIAEQTNNMLALNASTEAAARAGAGGFAVVAD 360
Db 301 EATDGTAGVEQJGERAADVESYTGVIDIAEQTNNMLALNASTEAAARAGAGGFAVVAD 360
QY 361 EVKALAEESREGQSTRVEBELVEQMAETEETVDQLDEYNORIGGVEREEMETLOEITD 420
Db 361 EVKALAEESREGQSTRVEBELVEQMAETEETVDQLDEYNORIGGVEREEMETLOEITD 420
QY 421 AVEDDASGMEVSTATDEQAVSTEEVAEMVDGVDRAGETAAALDDIADTDQOVRTEE 480
Db 421 AVEDDASGMEVSTATDEQAVSTEEVAEMVDGVDRAGETAAALDDIADTDQOVRTEE 480
QY 481 VRETGKLS 489
Db 481 VRETGKLS 489

```

RESULT 2

HMT_HALN1

ID HMT_HALN1

AC 09HPR6;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DE Heme-based aerotactic transducer hemAT.

GN HEMAT OR HTR10 OR VNG1505G.

OS Halobacterium sp. (strain NRC-1).

OC Archaea: Euryarchaeota: Halobacteriales;

OC Halobacteriaceae: Halobacterium.

OX NCBI_TaxID=64091;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20504483; PubMed=11016950;

RA Ng W.V., Kennedy S.P., Mahaltras G.G., Bergquist B., Pan M.,

RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Shrogha J.,

RA Swartzell S., Welt D., Hall J., Dahl T.A., Welt R., Goo Y.A.,

RA Lelthausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,

RA Madocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,

RA Isebnberger T.A., Beck R.F., Pohlischer M., Spidlich J.L., Jung K.-H.,

RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,

RA "Ehhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.,

RT Genome sequence of Halobacterium species NRC-1."

RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).

CC -1- FUNCTION: Heme-containing signal transducer responsible for

CC aerotaxis, the migratory response toward or away from oxygen (By

CC similarity).

CC -1- SUBUNIT: Homotetramer (Probable).

CC -1- SIMILARITY: BELONGS TO THE CHEMOTAXIS SENSORY TRANSDUCERS

CC FAMILY.

CC

CC

CC

CC

CC

CC

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL: AE005064; AAC19801.1; -.
DR HSSP: P02942; 1Q07.
DR InterPro: IP0004089; Chmtaxis_transd.
DR InterPro: IP0004090; Me_Chemotaxis.
DR Pfam: PF00015; MCPsignal.1.
DR PRINTS: PR00260; CHEMTRNSDUCR.
DR SMART: SM00283; MA.1.
KW Transducer; Heme; Complete proteome.
SQ SEQUENCE 489 AA; 52852 MW; 8C0AEF179667791E CRC64;

Query Match          98.6%; Score 2360; DB 1; Length 489;
Best Local Similarity 99.0%; Pred. No. 2e-95;
Matches 484; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MSNDNDTLVTADVRNGIDGHALADRIGLDEAEIAMRLSFTGIDDDPMALAAEQPLEFAT 60
Db 1 MSNDNDTLVTADVRNGIDGHALADRIGLDEAEIAMRLSFTGIDDDPMALAAEQPLEFAT 60
QY 61 ADALVTDFDHLSEYERTQDLPFANSTKTYBQLEKTAQAEYLLGREGYDTEYAQARATG 120
Db 61 ADALVTDFDHLSEYERTQDLPFANSTKTYBQLEKTAQAEYLLGREGYDTEYAQARATG 120
QY 121 KIHVDVGLGPDVYLGAATRYTGTGLDALADDDVADGEEAAAANDELVARFLPMLKLTFF 180
Db 121 KIHVDVGLGPDVYLGAATRYTGTGLDALADDDVADGEEAAAANDELVARFLPMLKLTFF 180
QY 181 DQOIAMDVTYIDSYAQRLLHDEIDSRQELANAVATHVAPLSLEATSDQVAERTDTMARAT 240
Db 181 DQOIAMDVTYIDSYAQRLLHDEIDSRQELANAVATHVAPLSLEATSDQVAERTDTMARAT 240
QY 241 DQOVDNRADVSREISSVSASVEEVAATADVRRTSDEAEALAOOGGEAADALATMTDID 300
Db 241 DQOVDNRADVSREISSVSASVEEVAATADVRRTSDEAEALAOOGGEAADALATMTDID 300
QY 301 EATDGTAGVEQJGERAADVESYTGVIDIAEQTNNMLALNASTEAAARAGAGGFAVVAD 360
Db 301 EATDGTAGVEQJGERAADVESYTGVIDIAEQTNNMLALNASTEAAARAGAGGFAVVAD 360
QY 361 EVKALAEESREGQSTRVEBELVEQMAETEETVDQLDEYNORIGGVEREEMETLOEITD 420
Db 361 EVKALAEESREGQSTRVEBELVEQMAETEETVDQLDEYNORIGGVEREEMETLOEITD 420
QY 421 AVEDDASGMEVSTATDEQAVSTEEVAEMVDGVDRAGETAAALDDIADTDQOVRTEE 480
Db 421 AVEDDASGMEVSTATDEQAVSTEEVAEMVDGVDRAGETAAALDDIADTDQOVRTEE 480
QY 481 VRETGKLS 489
Db 481 VRETGKLS 489

```

RESULT 3

HTR3_HALN1

ID HTR3_HALN1

AC 09H000;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE Halobacterial transducer protein III.

GN HTR9 OR VNG1395G.

OS Halobacterium sp. (strain NRC-1).

OC Archaea: Euryarchaeota: Halobacteriales;

OC Halobacteriaceae: Halobacterium.

OX NCBI_TaxID=64091;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20504483; PubMed=11016950;

RA Ng W.V., Kennedy S.P., Mahaltras G.G., Bergquist B., Pan M.,

RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Shrogha J.,

RA Swartzearl S., Mellr D., Hall J., Dahl T.A., Melti R., Goo Y.A.,
RA Lettner B., Keller K., Cruz R., Danson M.J., Hough C.M., Dale H.,
RA Maddocks D.G., Jablonksi P.E., Krebs M.P., Angevine D.W., Dale H.,
RA Isenbarger T.A., Peck R.F., Polischroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Eberhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.,
RA "Genome sequence of Halobacterium species NRC-1",
RT proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
RL
CC -1- FUNCTION: POTENTIALLY INVOLVED IN CHEMO- OR PHOTOTACTIC SIGNAL
CC TRANSDUCTION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE CHEMOTAXIS SENSORY TRANSDUCERS FAMILY.
CC -1- SIMILARITY: CONTAINS 1 PAS (PER-ARNT-SIM) DIMERIZATION DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE005058; AAC19717.1; -
DR HSSP: P02942; 10U7.
DR InterPro: IPR004089; Chmtaxis_transd.
DR InterPro: IPR003660; HAMF.
DR InterPro: IPR004090; Me_chemotaxis.
DR InterPro: IPR000014; PAS_domain.
DR Pfam: PF00015; MCPsignal; 1.
DR PRINTS: PR00260; CHEMTNSDUCR.
DR SMART: SM00304; HAMF; 1.
DR SMART: SM00283; MA; 1.
DR SMART: SM00091; PAS; 1.
DR PROSITE: PS50112; PAS; 1.
DR Transducer: Complete proteome.
FT DOMAIN 10
FT FT SEQUENCE 481 AA: 50773 MW: B333B5B5DFE206D1 CXC64;
SO
Query Match 26.9%; Score 643; DB 1; Length 481;
Best Local Similarity 37.6%; Pred. No. 1.8e-21;
Matches 182; Conservative 72; Mismatches 196; Indels 34; Gaps 10;
OY 15 NCIDGHALADRLGDLDEAELARLSTGTGDDTMAALAEOLPEFETADAL-VTFPYDHLF 73
DB 19 NTLDPAPAA-VDAAGAVAW-----DDQIAL-----LETPAEALIGVTDIGERLN 63
OY 74 SYERTQDFANSTKVEQDLKETQAYLILGEGVEDTEAQAQRATIGKIHVLGLGPGVY 133
DB 64 D-----DGRBALANKKADPIDDAHNEY---DGLADESTALITGDYVEEDTTVACNTDLM 116
OY 134 LGATRYRYTGLDALADVDVADRGEEAAAVDELAREFLPMKLK-----TFDQOI--- 184
DB 117 FIATPVYHTGREGVY-ELVQDR-SSARQYSEIQALGELVDLTIDAYDAGFDATVDA 174
OY 185 AMDYIYDSYQRLHDEIDSRQELANAVATHVEAPLSSLEPATISQDYAERTDTMRARTDOV 244
DB 175 AEDTLIDDEYIQIGNLTFEGDTLAAHTEHVNDRLEAASQAVSESSAEIDELSTAQ 234
OY 245 DRMAVRSREISSVSEVAFTADVDRTSEDAFALNOOGGAANDDLATTTDIDEND 304
DB 235 TIVSVATVEVELSTAVELSTADEVVDTSATFAERLDDGSAASDAADMADVATAD 294
OY 305 GVTAGVEQLGERAAVDESVTGVIDIASQTNMLALANASIEAARAGEGEFAVYADEKA 364
DB 295 SVTSQVEALQNRNIEDIDEVDYITGIAQTNMLALANASIEAARAGEBEGFAVYAEVKA 354
OY 365 LAESREOSTRVEELVEQMAETFEETVQDLDEVNORIGEGYERVEAMETLOETDAVED 424
DB 355 LAEDQASNAHGHTESLYSEIORDTAVTDVLTDTTRIEDIADVAQVEDAMASFEIYATVEA 414
OY 425 AASGQEVSTADDEAVSVEEYAVGVCDVRAGISAAALDDIADATQOQVTVVEVET 484
DB 415 TAEGLQVSDATNEQASHEETAAAVDETADLADDTITTAADITYSQTASAMLDHDES 474

Y	485	VGKL	488
Db	475	VSEL	478
Y	185	AMDTYIDSYAORLHDEIDSRQELANAVATHEAPLSSLEATISQDYAERTDTMRARTDOV	244
Y	117	FIATPVYHTGFRGI-EIVODR-SSSARYSGEALALGELVLDIATVADAGFADATVDA	174
Y	134	LGATRYTYTGLDLADLVADVADRGEEAAAVDELATFRLPMKLK-----TFDOI---	184
Y	117	FIATPVYHTGFRGI-EIVODR-SSSARYSGEALALGELVLDIATVADAGFADATVDA	174
Y	185	AMDTYIDSYAORLHDEIDSRQELANAVATHEAPLSSLEATISQDYAERTDTMRARTDOV	244
Y	15	NGIDGHALDRIGLDEAEIANRLSFTGIDDTNAAALAEQPLFEATADAL-VTDFYDHL	73
Y	19	NTLVDPAPFA--VDADGAVVAVW-----DDQIAL-----LETAPEDAIGVTIDGERLN	63
Y	74	SYERTQDLEFANSTVVEDEKLETOAEYLLGSLGRGEVDTEYTAQAQRATIGKHDVIGLGPVY	133
Y	64	D-----DGSRIALANKYADPPIDAHHEV---DVGGLADESTALLTGDTVEYEDTTVAGNTDLM	116
Y	134	LGATRYTYTGLDLADLVADVADRGEEAAAVDELATFRLPMKLK-----TFDOI---	184
Y	117	FIATPVYHTGFRGI-EIVODR-SSSARYSGEALALGELVLDIATVADAGFADATVDA	174
Y	185	AMDTYIDSYAORLHDEIDSRQELANAVATHEAPLSSLEATISQDYAERTDTMRARTDOV	244
Y	15	NGIDGHALDRIGLDEAEIANRLSFTGIDDTNAAALAEQPLFEATADAL-VTDFYDHL	73
Y	19	NTLVDPAPFA--VDADGAVVAVW-----DDQIAL-----LETAPEDAIGVTIDGERLN	63
Y	74	SYERTQDLEFANSTVVEDEKLETOAEYLLGSLGRGEVDTEYTAQAQRATIGKHDVIGLGPVY	133
Y	64	D-----DGSRIALANKYADPPIDAHHEV---DVGGLADESTALLTGDTVEYEDTTVAGNTDLM	116
Y	134	LGATRYTYTGLDLADLVADVADRGEEAAAVDELATFRLPMKLK-----TFDOI---	184
Y	117	FIATPVYHTGFRGI-EIVODR-SSSARYSGEALALGELVLDIATVADAGFADATVDA	174
Y	185	AMDTYIDSYAORLHDEIDSRQELANAVATHEAPLSSLEATISQDYAERTDTMRARTDOV	244
Y	15	NGIDGHALDRIGLDEAEIANRLSFTGIDDTNAAALAEQPLFEATADAL-VTDFYDHL	73
Y	19	NTLVDPAPFA--VDADGAVVAVW-----DDQIAL-----LETAPEDAIGVTIDGERLN	63
Y	74	SYERTQDLEFANSTVVEDEKLETOAEYLLGSLGRGEVDTEYTAQAQRATIGKHDVIGLGPVY	133
Y	64	D-----DGSRIALANKYADPPIDAHHEV---DVGGLADESTALLTGDTVEYEDTTVAGNTDLM	116
Y	134	LGATRYTYTGLDLADLVADVADRGEEAAAVDELATFRLPMKLK-----TFDOI---	184
Y	117	FIATPVYHTGFRGI-EIVODR-SSSARYSGEALALGELVLDIATVADAGFADATVDA	174
Y	185	AMDTYIDSYAORLHDEIDSRQELANAVATHEAPLSSLEATISQDYAERTDTMRARTDOV	244
Y	15	NGIDGHALDRIGLDEAEIANRLSFTGIDDTNAAALAEQPLFEATADAL-VTDFYDHL	73
Y	19	NTLVDPAPFA--VDADGAVVAVW-----DDQIAL-----LETAPEDAIGVTIDGERLN	63
Y	74	SYERTQDLEFANSTVVEDEKLETOAEYLLGSLGRGEVDTEYTAQAQRATIGKHDVIGLGPVY	133
Y	64	D-----DGSRIALANKYADPPIDAHHEV---DVGGLADESTALLTGDTVEYEDTTVAGNTDLM	116
Y	134	LGATRYTYTGLDLADLVADVADRGEEAAAVDELATFRLPMKLK-----TFDOI---	184
Y	117	FIATPVYHTGFRGI-EIVODR-SSSARYSGEALALGELVLDIATVADAGFADATVDA	174
Y	185	AMDTYIDSYAORLHDEIDSRQELANAVATHEAPLSSLEATISQDYAERTDTMRARTDOV	244
Y	15	NGIDGHALDRIGLDEAEIANRLSFTGIDDTNAAALAEQPLFEATADAL-VTDFYDHL	73
Y	19	NTLVDPAPFA--VDADGAVVAVW-----DDQIAL-----LETAPEDAIGVTIDGERLN	63
Y	74	SYERTQDLEFANSTVVEDEKLETOAEYLLGSLGRGEVDTEYTAQAQRATIGKHDVIGLGPVY	133
Y	64	D-----DGSRIALANKYADPPIDAHHEV---DVGGLADESTALLTGDTVEYEDTTVAGNTDLM	116
Y	134	LGATRYTYTGLDLADLVADVADRGEEAAAVDELATFRLPMKLK-----TFDOI---	184
Y	117	FIATPVYHTGFRGI-EIVODR-SSSARYSGEALALGELVLDIATVADAGFADATVDA	174
Y	185	AMDTYIDSYAORLHDEIDSRQELANAVATHEAPLSSLEATISQDYAERTDTMRARTDOV	244
Y	15	NGIDGHALDRIGLDEAEIANRLSFTGIDDTNAAALAEQPLFEATADAL-VTDFYDHL	73
Y	19	NTLVDPAPFA--VDADGAVVAVW-----DDQIAL-----LETAPEDAIGVTIDGERLN	63
Y	74	SYERTQDLEFANSTVVEDEKLETOAEYLLGSLGRGEVDTEYTAQAQRATIGKHDVIGLGPVY	133
Y	64	D-----DGSRIALANKYADPPIDAHHEV---DVGGLADESTALLTGDTVEYEDTTVAGNTDLM	116
Y	134	LGATRYTYTGLDLADLVADVADRGEEAAAVDELATFRLPMKLK-----TFDOI---	184
Y	117	FIATPVYHTGFRGI-EIVODR-SSSARYSGEALALGELVLDIATVADAGFADATVDA	174
Y	185	AMDTYIDSYAORLHDEIDSRQELANAVATHEAPLSSLEATISQDYAERTDTMRARTDOV	244
Y	15	NGIDGHALDRIGLDEAEIANRLSFTGIDDTNAAALAEQPLFEATADAL-VTDFYDHL	73
Y	19	NTLVDPAPFA--VDADGAVVAVW-----DDQIAL-----LETAPEDAIGVTIDGERLN	63
Y	74	SYERTQDLEFANSTVVEDEKLETOAEYLLGSLGRGEVDTEYTAQAQRATIGKHDVIGLGPVY	133
Y	64	D-----DGSRIALANKYADPPIDAHHEV---DVGGLADESTALLTGDTVEYEDTTVAGNTDLM	116
Y	134	LGATRYTYTGLDLADLVADVADRGEEAAAVDELATFRLPMKLK-----TFDOI---	184
Y	117	FIATPVYHTGFRGI-EIVODR-SSSARYSGEALALGELVLDIATVADAGFADATVDA	174
Y	185	AMDTYIDSYAORLHDEIDSRQELANAVATHEAPLSSLEATISQDYAERTDTMRARTDOV	244
Y	15	NGIDGHALDRIGLDEAEIANRLSFTGIDDTNAAALAEQPLFEATADAL-VTDFYDHL	73
Y	19	NTLVDPAPFA--VDADGAVVAVW-----DDQIAL-----LETAPEDAIGVTIDGERLN	63
Y	74	SYERTQDLEFANSTVVEDEKLETOAEYLLGSLGRGEVDTEYTAQAQRATIGKHDVIGLGPVY	133
Y	64	D-----DGSRIALANKYADPPIDAHHEV---DVGGLADESTALLTGDTVEYEDTTVAGNTDLM	116
Y	134	LGATRYTYTGLDLADLVADVADRGEEAAAVDELATFRLPMKLK-----TFDOI---	184
Y	117	FIATPVYHTGFRGI-EIVODR-SSSARYSGEALALGELVLDIATVADAGFADATVDA	174
Y	185	AMDTYIDSYAORLHDEIDSRQELANAVATHEAPLSSLEATISQDYAERTDTMRARTDOV	244
Y	15	NGIDGHALDRIGLDEAEIANRLSFTGIDDTNAAALAEQPLFEATADAL-VTDFYDHL	73
Y	19	NTLVDPAPFA--VDADGAVVAVW-----DDQIAL-----LETAPEDAIGVTIDGERLN	63
Y	74	SYERTQDLE	

Db 175 AEDTLDDERYIGRNLTEFGDTLLAHTEVHNDVERLEAASQAVSSSESEIDELSTAQS 234
 QY 245 DRADVSREISSVSASVEEVAASTADVRRSEDAEALAOGEAADAALATMTDIDPATD 304
 Db 235 TNSVTATEVETLSATQELASTADEVDTSATRELRADGSAASASADAMADVATRAD 294
 QY 305 GVTAGVEQLGERADAVESVTVGIDIDIAEQTNMLALNLSIEAARAGEGFAVVADEVKA 364
 Db 295 SVTSDVEALNRLEDIEDVVDITGIEQTNMLALNLSIEAARAGEGFAVVADEVKA 354
 QY 365 LAESRQSRQSRVEELVQOMAEETFEVDOLDEYNORIGGEVEREEMNELOETITDVED 424
 Db 355 LAEDQSNMAGHIESVSEIRDADVTDTLTTRIEDVAVQVEDMASFEETVATVEA 414
 QY 425 AASGMOEVSTATDEQAVSTEVAEMVGVDDRAGEIAALDIDATDQCVRTVEEVRET 484
 Db 415 TAEIGEVQATNEQOASAEIAAMVDETADLDDITTAADIVSQHEASTLIXDDES 474
 QY 485 VGKL 488
 Db 475 VSEL 478
 RESULT 5
 HTR6_HALN1 STANDARD; PRT: 778 AA.
 ID HTR6_HALN1
 AC 09HR92;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Halobacterium transducer protein VI.
 GN HTR4 OR HTPVI OR VNG0806G.
 OS Halobacterium sp. (strain NRC-1).
 OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;
 OC Halobacteriales; Halobacterium.
 OX NCBI_Taxid=64091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20504483; PubMed=11016950;
 RA N.W.V., Kennedy S.P., Mahilras G.G., Bergquist B., Pan M.,
 RA Shukla H.D., Lasky S.R., Ballas N.S., Thorsson V., Shrogha J.,
 RA Swartzell S., Welt D., Hall J., Dahl T.A., Melit R., Goo Y.A.,
 RA Leitner B., Keller K., Cruz R., Danson M.J., Hough D.W.,
 RA Maddocks D.G., Jadhavski P.E., Krebs M.P., Angevine C.M., Dale H.,
 RA Isenbarger T.A., Beck R.F., Pohlischer C.J., Spudis J.L., Jung K.-H.,
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
 RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
 RA Genome sequence of Halobacterium species NRC-1";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
 CC - FUNCTION: POTENTIALLY INVOLVED IN CHEMO- OR PHOTOTACTIC SIGNAL
 CC TRANSDUCTION (BY SIMILARITY).
 CC - SIMILARITY: BELONGS TO THE CHEMOTAXIS SENSORY TRANSDUCERS FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AE005022; AG19266.1; -.
 DR HSSP: P02942; 1007.
 DR InterPro: IPR004089; Chmtaxis_transd.
 DR InterPro: IPR003660; HAM.
 DR InterPro: IPR004080; Me_Chemotaxis.
 DR Pfam: PF00015; MCPsignal.1.
 DR Pfam: PF00672; HAM.1.
 DR PRINTS: PR00260; CHEMTRNSDUC.
 DR SMART: SM00304; HAM.2.
 DR SMART: SM00283; MA.1.
 KM Transducer; Transmembrane; Complete proteome.
 FT DOMAIN 1 26 CITOPASMIC (POTENTIAL).

FT TRANSMEM 27 47 POTENTIAL.
 FT DOMAIN 48 296 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 297 317 POTENTIAL.
 FT DOMAIN 318 778 CYTOSPLASMIC (POTENTIAL).
 SQ SEQUENCE 778 AA; 82076 MW; D752278727A4FA79 CRC64;
 Query Match 24.1%; Score 578; DB 1; Length 778;
 Best Local Similarity 33.7%; Pred. No. 1.9e-18;
 Matches 145; Conservative 75; Mismatches 172; Indels 38; Gaps 6;
 QY 88 TVGQLKETQAEYLLGRCGYDTEFAQAQR-RIGKIHVDVGLGPPDYLLCATRYTGILD 146
 Db 320 TVRALNDLEAK-AAALERGEYDIDLVARVDELGRLEFAEPASLDYVQARIRDAEQVD 378
 QY 147 -----ALADVVADGEEAQAADVLVRF-----PMLKLTFDQOIAMD 187
 Db 379 AEAARSEAEQAQDAEAEQAQAEAREESQAQRLETTAEARSEIMRAYAGDLVRLD 438
 QY 188 TYIDSYAQ-----RLHDEISROELANAVATHEAPLSLEATSDQVAE 231
 Db 439 ADVEQAAMADIAAFNEMADMEATTADVVAFADEVATA-STDASQAQAAVEQIGRDVSD 497
 QY 232 RTDTMARATDQVDRADVSREISSVSASVEEVAASTADVRRSEDAEALAOGEAQAAD 291
 Db 498 AVGRIRDRADQDRDQLEVAASETDEMSATLEEVAAASADQVAETSORAAALGDGQAAD 557
 QY 292 ALATMTDIDEATDGTAGVGEQLGERADAVESVTVGIDIDIAEQTNMLALNLSIEAARAGEA 351
 Db 558 AVAQLEIEDEFTQAATRAYDDELAKMSEIFTYAALTDIAEQTNMLALNLSIEAARADQ 617
 QY 352 GEGFAVVADEKVALAESRQSTRVELVEQOMAEETFEVDOLDEYNORIGGEVEREVEA 411
 Db 618 GDFEFAVADEKVALDESKASAEIEBALVEBAQETSVAAADRIOERVSDEVEVSET 677
 QY 412 METLOETTAVDEAASGMOEVSTATDEQAVSTEVAEMVGVDDRAGEIAALDIDATD 471
 Db 678 ERSLSEIAGRIADPTGVQELSNAMDQAASVDTTAVGDAVALGEETATEAESTRAD 737
 QY 472 DQCVRTVEEV 481
 Db 738 AEGATTLSDV 747
 RESULT 6
 HTR6_HALSA STANDARD; PRT: 778 AA.
 ID HTR6_HALSA
 AC 048319;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Halobacterium transducer protein VI.
 GN HTR4 OR HTPVI.
 OS Halobacterium salinarum.
 OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;
 OC Halobacteriales; Halobacterium.
 OX NCBI_Taxid=2242;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=59;
 RX MEDLINE=96275896; PubMed=8674984;
 RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,
 RA Oesterhelt D.;
 RT "A family of halobacterial transducer proteins";
 RL FEWS Microbiol. Lett. 139:161-168(1996).
 CC - FUNCTION: POTENTIALLY INVOLVED IN CHEMO- OR PHOTOTACTIC SIGNAL
 CC TRANSDUCTION.
 CC - SIMILARITY: BELONGS TO THE CHEMOTAXIS SENSORY TRANSDUCERS FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch)

CC -----

DR EMBL: X95590; CA64843.1; -

DR HSSP: P02942; 1007.

DR InterPro: IPR004089; Cmltaxis_transd.

DR InterPro: IPR003660; HAMF.

DR Pfam: PF00015; MCPsignal; 1.

DR Pfam: PF00672; HAMF; 1.

DR SMART: SM00304; HAMF; 2.

DR SMART: SM00283; MA; 1.

DR Transducer: Transmembrane.

FT DOMAIN 1 26 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 27 47 POTENTIAL.

FT DOMAIN 48 296 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 297 317 POTENTIAL.

FT DOMAIN 318 778 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 318 778 CYTOPLASMIC (POTENTIAL).

SQ SEQUENCE 778 AA; 82077 MM; 134CD7F0A3334CD CRC64;

Query Match 23.9%; Score 573; DB 1; Length 778;

Best Local Similarity 33.0%; Pred. No. 3.1e-18;

Matches 144; Conservative 74; Mismatches 168; Indels 50; Gaps 6;

OY 88 TVEOLKETQAEYLGLGGEYDTEYAQRA-RIGKIHVILGIPVYLGATRYTGLLD 146

DB 320 TYRALNDLEAK-AALERGEYDTLDVARVDELGRFLFEAFSLRTVGARIDANEQVD 378

OY 147 -----ALADVVADRGEEAANADELVARFL-----PMLKLTFFDOQIAMD 187

DB 379 AEAARSEAAQAQDAEAAQAEAEAREESEARLETTAEAFSETRMAYAGDLTVRAD 438

OY 168 TYIDSYAO-----RLHDEIDSQELANNAVTHVEAPLSLEAT 225

DB 439 ADVEAQAADIAAFNEMAADKEATIDAGRADEVAT-----ASTDASAAAEQT 491

OY 226 SODVARTDTMRARTDDQVDRMADVSREISSVSASVEEYASTADVRRTSSEDAEALAOGE 285

DB 432 GNDVSDANGRIADRRAADQDQLENAVSEFDEMSATIEEVAASADQVAETSSQAAALGDDG 551

OY 286 EAADALATMTDIDEATDGTAGVEQLGERADYESTGTVIDDIAEQTMALANSTEA 345

DB 552 QAADDAVAQLEIEDETCQAATAVDLEAKMSEIETIYAATTDIAEQTMALANSTEA 611

OY 346 AAGEGEGFAVVADEVKALAEESRQSRVELEVGOMAEETEVYDQLEVNORIGEV 405

DB 612 ARAADGDEFAVVADEVKALAEESRQSRVELEVGOMAEETEVYDQLEVNORIGEV 671

OY 406 EREVEAMETLOETDAVEDAASGMOEVSTATDEQAVSTEVEVEMVGVDRAGEIAALD 465

DB 672 ETVSETERLSLEIAGRIAEADTGVQEISNAMDDQAAVSVDVTTVGVAAALGEFTATEAE 731

OY 466 DIADATDOOVRTVEEV 481

DB 732 STADAAAEQASTLSDV 747

RESULT 7

HTF4_HALN1 STANDARD: PRT; 810 AA.

ID HTF4_HALN1

AC Q9HP84; 16-OCT-2001 (Rel. 40; Created)

DT 16-OCT-2001 (Rel. 40; Last sequence update)

DT 16-OCT-2001 (Rel. 40; Last annotation update)

DE Halobacterium sp. (strain NRC-1).

GN HTF4 OR HTF4 OR VNG1760G.

OS Halobacterium sp. (strain NRC-1).

OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;

CC Halobacteriaceae; Halobacterium.

NCBI_TaxID=64091;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20504483; PubMed=11016950;

RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,

RA Shukla H.D., Lasky S.R., Ballig N.S., Thorsson V., Spirogna J.,

RA Swartwell S., Well T.A., Hall J., Dahl T.A., Welti R., Goo Y.A.,

RA Leithausen B., Keller K., Cruz R., Danon M.J., Hough D.W.,

RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,

RA Isenbarger T.A., Beck R.F., Pohlischroder M., Spudis J.L., Jung K.-H.,

RA Alam M., Freltas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,

RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.,

RA "Genome sequence of Halobacterium species NRC-1."

RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).

CC -1- FUNCTION: POTENTIALLY INVOLVED IN CHEMO- OR PHOTOTACTIC SIGNAL

CC -1- TRANSDUCTION (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE CHEMOTAXIS SENSORY TRANSDUCERS FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch)

CC -----

CC EMBL: AE005080; AAG19986.1; -

DR HSSP: P02942; 1007.

DR InterPro: IPR004089; Cmltaxis_transd.

DR InterPro: IPR003660; HAMF.

DR Pfam: PF00015; MCPsignal; 1.

DR Pfam: PF00672; HAMF; 1.

DR SMART: SM00304; HAMF; 2.

DR SMART: SM00283; MA; 1.

DR Transducer: Transmembrane; Complete proteome.

FT DOMAIN 1 38 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 39 59 POTENTIAL.

FT DOMAIN 60 323 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 324 344 POTENTIAL.

FT DOMAIN 345 810 CYTOPLASMIC (POTENTIAL).

SQ SEQUENCE 810 AA; 85219 MM; 4BF3648BD228BD80 CRC64;

Query Match 23.2%; Score 555; DB 1; Length 810;

Best Local Similarity 32.1%; Pred. No. 1.9e-17;

Matches 142; Conservative 85; Mismatches 174; Indels 42; Gaps 8;

OY 84 NSTKVEOLKETQAEYLGLGGEYDTEYAQRA-RIGKIHVILGIPVYLG 135

DB 346 NTAANVQSLAAAE-----TEAGNDVYASSHROEIGQLFASISMRALATQIDEAEA 401

OY 136 ATTRYTGLDLAD---DVADGEEAANADELVA-----RFLPMLKL----- 178

DB 402 AREQATEAQDAEAREERAEADRAEDAKADFEALAELEQAERYSDVMAACADGILT 461

OY 179 -----FPDQILA--MDYIDSYAORLHDEIDSQELANNAVAT---HYEAPLSLEATS 226

DB 462 RRPADDDTDNEAANAIAASFNEMLAOWEHTIITL-DEPADAVATVSEEEVGAADAEARS 520

OY 227 QDVARTDTMRARTDDQVDRMADVSREISSVSASVEEYASTADVRRTSSEDAEALAOGE 286

DB 521 GGVSESVQEIACADEQRNMLDTVSGEMTDLFAIEEVAASADSVAEHSQPAEIAARDGE 580

OY 287 AAADALATMTDIDEATDGTAGVEQLGERADYESTGTVIDDIAEQTMALANSTEA 346

DB 581 QTAEDAIERSLTVQEAIDATVONVEALDDOMARISEIYDLIDIAEQTMALANSTEA 640

OY 347 RAGEGEGFAVVADEVKALAEESRQSRVELEVGOMAEETEVYDQLEVNORIGEV 406

DB 641 RADKSGDGFVAVVADEVKALAEETQSAGCIERHITVEVQSOTATVAEAAAEASPDAGID 700

OY 407 REVEAMETLOETDAVEDAASGMOEVSTATDEQAVSTEVEVEMVGVDRAGEIAALD 466

DB 701 AVEEVYDAFTAVSDAHADETDGTGVQEISNAMDDQAAVSVDVTTVGVAAALGEFTATEAE 760

OY 467 IADATDOOVRTVEEVRYGKLS 489

DB 761 VSAAAEEQAASMSISDSVESLS 783

```

RESULT 8
HTR4_HALSA
ID HTR4_HALSA STANDARD: PRT: 810 AA.
AC 048317:
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Halobacterial transducer protein IV.
GN HTR5 OR HTRIV.
OS Halobacterium salinarum.
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
ON NCBI_TaxID=2242;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=59;
RX MEDLINE=96275896; PubMed=8674984;
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,
RA Oesterhalt D.;
RT "A family of halobacterial transducer proteins.";
RL FEMS Microbiol. Lett. 139:161-168(1996).
CC -1- FUNCTION: POTENTIALLY INVOLVED IN CHEMO- OR PHOTOTACTIC SIGNAL
CC TRANSDUCTION.
CC -1- SIMILARITY: BELONGS TO THE CHEMOTAXIS SENSORY TRANSDUCERS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X95589; CAA64841.1; -
DR HSSP: P02942; 1OU7.
DR InterPro: IPR004089; Chmtaxis_transd.
DR InterPro: IPR003660; HAMP.
DR Pfam: PF00015; MCPsignal; 1.
DR Pfam: PF00672; HAMP; 1.
DR SMART: SM00304; HAMP; 2.
DR SMART: SM00283; MA; 1.
KW Transducer; Transmembrane.
FT DOMAIN 1 38 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 39 59 POTENTIAL.
FT DOMAIN 60 323 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 324 344 POTENTIAL.
FT DOMAIN 345 810 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 810 AA; 85207 MW; 3780F6046A39D9BA CRC64;

Query Match 23.1%; Score 554; DB 1; Length 810;
Best Local Similarity 32.1%; Pred. No. 2, 1e-17;
Matches 142; Conservative 85; Mismatches 174; Indels 42; Gaps 8;

OY 84 NSTKTYEQKETOAEYLIGRGEYDEYTAQR-----ARIGKHVYLGLGPPVYLG 135
DB 346 NTAQAQSSAAMAE-----IEAGNYDVVASRDEIGLGFASIGSRDLALVQIDEAEA 401
OY 136 AYTRYTYGLDLAD-----DVADRGEBAAMAADELVA-----PRLPLKLL----- 178
DB 402 AREQATEAQDQDAEREREDARERADAKADAEALAELEAQAERYSDVMAACADGDLT 461
OY 179 -----TFDQQLA--MDYIYSVAORLHDEISROELANAVT---HVEAPLSLEATS 226
DB 462 RRMADDTDNEAMATASNEMLAOMEHTIIDI-QEFDAYVATASEEAIVEGAADPERAS 520
OY 227 QDVAERTDTMRATDDQVDMADYSREISSVASVEEVASTADVARTSEDAALAQGE 286
DB 521 GQVSESEVQEIAGAADQRNRLDTVSGEMTDLSAIEEVAASADSVASHSHQTEIARDGE 580
OY 287 AAMDALATMTDIDEDATDGTAGVGEOLGERADYVESVTGVIDIAQTNLALNASTIEA 346
DB 581 QTRDAIERSLSVQEAIDATVONVEALDDQMAEISSEIVLDSIAQTNLALNANIEEA 640

```

```

OY 347 RAGEAGEGFVAVDADEVKALAEESRQSTRVELVEQMAETEETVDOLDEVNQRIGGYE 406
DB 641 RADKSGDGFVAVDADEVKALAEETQESAGDIERRITEVQSQTATVAEARAESMDAGID 700
OY 407 RVEAMETLQETDVAEDDAASGQEVSTATDEQAVSTEVEAEVGDVDRAGITAAALDD 466
DB 701 AVEEVDAFTVAVDHDEDTGQVQISDPTDQASSTEEAVSMTEVADLSDSTAGEAQS 760
OY 467 IADATDOQVTEVEEVEVETGKLS 489
DB 761 VSAAEQQAASMEISDSVESTS 783

RESULT 9
HTR2_NATPH
ID HTR2_NATPH STANDARD: PRT: 534 AA.
AC P42259;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sensory rhodopsin II transducer (HTR-II) (Methyl-accepting phototaxis
DE protein II) (MPP-II).
GN HTR2 OR HTRII.
OS Naeromonas pharaonis (Naerobacterium pharaonis).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Naeromonas.
ON NCBI_TaxID=2257;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP-1 / 28;
RX MEDLINE=95224074; PubMed=7708770;
RA Seidel R., Scharf B., Gautel M., Knebel K., Oesterhalt D.,
RA Engelhard M.;
RT "The primary structure of sensory rhodopsin II: a member of an
RT additional retinal protein subgroup is coexpressed with its
RT transducer, the halobacterial transducer of rhodopsin II.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:3036-3040(1995).
CC -1- FUNCTION: TRANSDUCES SIGNALS FROM THE PHOTOTAXIS RECEPTOR
CC SENSORY RHODOPSIN II (SR-II) TO THE FLAGELLAR MOTOR. RESPONDS TO
CC LIGHT CHANGES THROUGH THE VARIATION OF THE LEVEL OF METHYLATION.
CC -1- SIMILARITY: BELONGS TO THE CHEMOTAXIS SENSORY TRANSDUCERS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z35086; CAA84468.1; -
DR HSSP: P02942; 1OU7.
DR InterPro: IPR004089; Chmtaxis_transd.
DR InterPro: IPR003660; HAMP.
DR Pfam: PF00015; MCPsignal; 1.
DR Pfam: PF00672; HAMP; 1.
DR SMART: SM00304; HAMP; 2.
DR SMART: SM00283; MA; 1.
KW Transducer; Photoreceptor; Transmembrane; Methylation.
FT DOMAIN 1 21 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 22 42 POTENTIAL.
FT DOMAIN 43 58 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 59 79 POTENTIAL.
FT DOMAIN 80 534 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 534 AA; 56622 MW; CBB801ZCCSF278E8 CRC64;

Query Match 22.5%; Score 537.5; DB 1; Length 534;
Best Local Similarity 31.2%; Pred. No. 7, 1e-17;
Matches 150; Conservative 95; Mismatches 177; Indels 59; Gaps 12;

OY 16 GIDGHALADRIQIDEA-----ETAWRLSTFGIDDTMAALAAEOPLEFATRADALVTOPY-- 69

```


Db 72 GINLGLVAAATLGGDTAASTLTAAKASRMG-DGDDLVEL-----ETFREDEIGDLVAA 123
 QY 70 -DHLESYERTODLFANSTVTBOLKETOAEYLLGGRGEBTDEYAAQRARIKHIVLGL 128
 Db 124 EDEMROSVTSLEDAKNAEDAEQAOKRAE-----BINFLOEAERFEVMDRCAD 175
 QY 129 GPDVYLGATRYTGLDALADVVADNREEAAAVDELVARFLPKLTFDQOIAMPT 188
 Db 176 GD-----FTQRLDAETDN-----EAMQSTEG-----SENM--MDG 204
 QY 189 YIDSTAQRLHDEIDSRKELANNAVATHEAPLSSLEATSDVAERTDTPARDDQVDRMA 248
 Db 205 -TEALVGRTERPDAVSEAEAEVRAAAE--SVMEA-SEVAVRAVONISDAAGDQETEVQ 259
 QY 249 DYSREISSASVSEEAASVADAVRRTSEDAEALAOQGEAALDALATMTDIDEATDGVTA 308
 Db 260 QTALEMDVDSATTEEAASADIAKTAARAQAEETGEGRTAETATTEMVESRFEQAVA 319
 QY 309 GVEQLGERAADVSVYGVDDIAEOTNMLNLSIFAARAGEGEGFAVVADEVKALAE 368
 Db 320 SMEBLNEDVREIJEVSEMIADIAEQTNILNLSIFAARADNSGFAVVADEVKALAE 379
 QY 369 SREQSTRVELVEQMOAETEEVVDOLDEVNQRIGSEVERVEEAMETLQETDAVEDAASG 428
 Db 380 TKAATEEIDDLIGTVODRTQTIVDDIRFSDQVSEGETVEDVDALERTIVDSVERTNDG 439
 QY 429 MEVSTATDEQAVSTEEVAVEMVDGVDNAGETAAALDDIADTDQVRYVEEYREYVGL 488
 Db 440 IDEINOSTDAQAADAAQKATTWEDMAATSEQTASDAETAEETEQAESVKEVEFDLIDL 499
 QY 489 S 489
 Db 500 S 500

RESULT 10

HTR5_HALN1
 ID HTR5_HALN1 STANDARD: PRT: 545 AA.
 AC 048318: 09HP85:
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Halobacterial transducer protein V.
 GN HTR7 OR HTPV OR VNG17596.
 OS Halobacterium sp. (strain NRC-1), and
 OS Halobacterium salinarium.
 OC Archaea: Euryarchaeota: Halobacteria: Halobacteriales:
 OC Halobacteriaceae: Halobacterium.
 OX NCBI_TaxID=64091, 2242;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRC-1;
 RX MEDLINE=20504483; PubMed=11016950;
 RX NG W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
 RA Shwela H.D., Laszlo S.R., Baliga N.S., Thorsson V., Schooga J.,
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Weir R., Goo Y.A.,
 RA Leitner B., Keller K., Cruz R., Danson M.J., Hough D.W.,
 RA Maddock D.G., Jablonksi P.E., Krebs M.P., Angevine C.M., Dale H.,
 RA Isenbacher T.A., Beck R.F., Pohlischer M., Spudich J.L., Jung K.-H.,
 RA Altm M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
 RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.,
 RT "Genome sequence of Halobacterium species NRC-1,"
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
 RU [2]
 RN SEQUENCE FROM N.A.
 RP SPECIES=H. salinarium; STRAIN=89;
 RX MEDLINE=96275896; PubMed=8674984;
 RX Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,
 RA Oesterhelt D.,
 RT "A family of halobacterial transducer proteins,"
 RL FEBS Microbiol. Lett. 139:161-168(1996).
 CC -1- FUNCTION: POTENTIALLY INVOLVED IN CHEMO- OR PHOTOTACTIC SIGNAL
 CC TRANSDUCTION.

CC -1- SIMILARITY: BELONGS TO THE CHEMOTAXIS SENSORY TRANSDUCERS FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC -----
 CC EMBL: AE005080; AAC19985.1; -
 CC DR EMBL: X95589; CA64842.1; -
 CC DR HSP: P02942; 1007.
 CC DR Interpro: IPR004089; Chmtaxis_transd.
 CC DR Interpro: IPR003660; HAMF.
 CC DR Pfam: PF00015; MCPsignal; 1.
 CC DR SMART: SM00304; HAMF; 1.
 CC DR SMART: SM00283; MA; 1.
 CC DR Transducer: Transmembrane; Complete proteome.
 CC FT TRANSMEM 10 30 POTENTIAL.
 CC FT TRANSMEM 44 64 POTENTIAL.
 CC FT TRANSMEM 80 100 POTENTIAL.
 CC SQ SEQUENCE 545 AA; 57070 MW; FDD870389C2F428B CRC64;
 Query Match 21.8%; Score 521; DB 1; Length 545;
 Best Local Similarity 29.6%; Pred No. 3.7e-16;
 Matches 142; Conservative 94; Mismatches 195; Indels 48; Gaps 10;
 QY 28 LDEAEIAMRLSFTGIDDDTMAALAEQPLEATADALVTFD---YDHLESYERTODLPA 83
 Db 64 LEKAGVATQI-FNLATGPMATTVAVFALIGTMTALIVEDMEALVERAQRQAEAEERA 122
 QY 84 NSTKYVEQLKETQAEYLLGLGGEYDTETAAOAFARIGKHIVGLGCPDVYLGATRYTIG 143
 Db 123 EAERAREKAEQKQAE---AEQTAESAESAKODAR-ERSAEIEQLAAD--LESQATEVGA 175
 QY 144 LLDALAD-----DVVADRG--EAAAVDELVAFLPLMLKLPFDQOIAMDTYIDSYA 194
 Db 176 TLEAASDGLTARVDATTNNAIEAYATVND-----LTTMERIT----- 216
 QY 195 QRLHDEIDSRKELANNAVATHEAPLSSLEATSDVAERTDTPAKARTDQVDRMAVSRRI 254
 Db 217 ----DEIQGFSTNVVTASREATAKAEIQDASCTVSSEVQELIAGGDDQREGLESVAEEM 272
 QY 255 SSVASVSEVASTADDVRTSEDAEALAOQGEAALDALATMTDIDEATDGVTAQVEQLG 314
 Db 273 DTSATVEVAVATASVATADTTDVAATAGQCTEDADIDAIDAOETQTVVANYDALE 332
 QY 315 ERAADVESYGVDDIAEOTNMLNLSIFAARAGEA-----GEGFAVVADEVKALAEER 370
 Db 333 DLTTEIDIAELISDAEOTNMLNLSIFAARAGSGGSGGSGGFAVVADEVKALAEER 392
 QY 371 EOSTREVELVEQMOAETEEVVDOLDEVNQRIGSEVERVEEAMETLQETDAVEDAASGQ 430
 Db 393 RSKADIAELIEEQSGTATVEIRAEGRVNDGAAVAEYDAPGAVENIENQETDGVQ 452
 QY 431 EVSTATDEQAVSTEEVAVEMVDGVDNAGETAAALDDIADTDQVRYVEEYREYVGL 489
 Db 453 EISQAMDEQAORSERVSSVDIATISQATDAERVAENVAASEQTASITVEVTSLSLSLA 511
 RESULT 11
 HTR2_HALVA
 ID HTR2_HALVA STANDARD: PRT: 433 AA.
 AC P42258;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Sensory rhodopsin II transducer (HTR-II) (Methyl-accepting phototaxis
 DE protein II) (MP-II) (Fragment).
 GN HTR11.
 OS Halococcus vallismortis.
 OC Archaea: Euryarchaeota: Halobacteriales: Halobacteriales:

Dh 389 KDAQTARALBSAADYEALTAVALADGDLTRRVDSARDHDMARIGHALNDMLDIETSYV 448
Qy 206 ELANAVARTHVEAPLSSLEA-----TSQDAERTDTRARTDOQVDMAYNSREISSVS 258
Dh 449 AAATFSDHVSADMAORVEDAGDAIDACTDVSTAVDEISDCAETEOTDLHEVAGVEDLS 508
Qy 259 ASVEEASTADVDRTSDEAEALAOQGEAADDALATATIDIDEATDGTAGVEQGERAA 318
Dh 509 ASAEVATVASTADTQAQASAVVDGQATGEDAVETMDVADDAEAAADMDALDSEMA 568
Qy 319 DVESVTGVYDIAEOTNMLALNASTIARAGEGFAVVADEVKALAEESRQSTVVEE 378
Dh 569 DIGEIVDIADIDQTNMLALNASTIARAGADGCFVAVVADEVKTLAEESRDAEDIES 628
Qy 379 LVEQMAETEETVQDLDEVNQRIGGEVERBERAMETLOEITDAVDAASGMOEVSTATDE 438
Dh 629 RLALAQGVSDVADDMRTSDTFSVSGRAVTDGDAATALDDVVSFAVDTTAAEITAAADR 688
Qy 439 QAVSTEVEAEMVDGVDRAGEIAALDDIADATDQVFTVEEVEETVGLS 489
Dh 689 QAHASRVAASAVDEVAGISOETAAQATVADSAATQTDTLSSVDAAADLA 739

RESULT 13

HTRL_HALN1
ID HTRL_HALN1 STANDARD: PRT: 535 AA.

AC P33741: O9HPF6: 28, Created)

DT 01-FEB-1994 (Rel. 28, Created)

DT 15-JUN-1994 (Rel. 29, Last sequence update)

DE Sensory rhodopsin I transducer (HTR-1) (Methyl-accepting phototaxis protein I) (MCP-1).

OS Halobacterium sp. (strain NRC-1), and

OC Archaea: Euryarchaeota: Halobacteriales: Halobacteriaceae: Halobacterium.

OX NCHL_Texid=64091, 2242;

RN NCHL_Texid=64091, 2242;

RA MEDLINE-20504483; PubMed-11016950;

RA NG W.V., Kennedy S.P., Mahlras G.G., Bergquist B., Pan M.,

RA Shukla H.D., Leaky S.R., Baliga N.S., Thorson V., Spingola J.,

RA Swartzell S., Weir D., Hall J., Dahl T.A., Weir R., Goo Y.A.,

RA Leithausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,

RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angewine C.M., Dale H.,

RA Isenberger T.A., Peck R.F., Pohlischer M., Spudich J.L., Jung K.-H.,

RA Alm M., Freltas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,

RA Gehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarina S.,

RA Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).

RA Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).

RA Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).

RA Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).

RA Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).

RA Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).

RA Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).

RA Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).

DR EMBL: AE005075; AAC19913.1; -.
DR EMBL: L05603; AAA72315.1; -.
DR PIR: A47190; A47190.
DR HSSP: P02942; 1007.
DR InterPro: IPR004089; Chmtaxis_transd.
DR Pfam: PF00015; MCPsignal; 1.
DR SMART: SM00304; HAMP; 2.
DR SMART: SM00283; MA; 1.
DR Transducer: Photoreceptor; Transmembrane; Methylation;
KW Complete proteome.
FT DOMAIN 1 13 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 14 28 POTENTIAL.
FT DOMAIN 29 38 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 39 54 POTENTIAL.
FT DOMAIN 55 535 CYTOPLASMIC (POTENTIAL).
FT MOD_RES 265 265 METHYLATION.
FT MOD_RES 272 272 METHYLATION.
FT MOD_RES 279 279 METHYLATION.
FT MOD_RES 463 463 METHYLATION.
FT MOD_RES 472 472 METHYLATION.
SQ SEQUENCE 535 AA; 56544 MW; B9945E4F6A9D091 CRC64;

Query Match 19.4%; Score 463.5; DB 1; Length 535;
Best Local Similarity 27.5%; Pred. No. 1.1e-13;
Matches 147; Conservative 87; Mismatches 205; Indels 95; Gaps 15;

Qy 2 SUNDNLTAVDRNCIDHALLADRLGDEAEIAMLSTFGIDDDTMALAAEQPLFEATA 61

Dh 31 TNDVSTIVA-----GIAG--LLTGSINAEIYVSIKELAAOTEVANGNLEQVTSRT 84

Qy 62 DALVDFYDHLSEYERTODLEFANSKTYEQ--LEKTOAEYLLGLGGEYDTEYAAQRA 117

Dh 85 D-----EFGSLDSIRQMOSLRGLNEMERTRADLEETQAE--AETAREBAQAKQEA 136

Qy 118 RIGKHIDVGLGCPDYLYLGYTRYTGGLDALADVYADRG--EAAAAYDELVAFLPML 175

Dh 137 QAERE-----ARELAATYQDPAKKRYGEMEAATG----- 168

Qy 176 KLTTFDQOIAMDT-----YIDYVQRLHDEIDSQELANAVATHEAPLSSLEATSDV 229

Dh 169 ----LTQVNDVDTHEAETWGTAFANQMDLQAVRTVYADIDEAKTEKMSSETADI 224

Qy 220 -AERTDTMRA-----RTDDQVDRADVSREISVSASVEEYASTADVRTSDEAEALA 282

Dh 225 EASAGDTVEAVSKISQANDQRTLELSAADVQVQSASAEIATITDPLASREDEVATAS 284

Qy 283 QQGEAAADALATMTDIDEATDGVAGVEQGERADYVESYGVYDIDIAEOTNMLALNAS 342

Dh 285 DAARDSSKSLADEMSISITVEDVAVGOVLNDVYAEITDIDYDITDIDGEOINMLALNAS 344

Qy 343 IEARAG--EAGEGFAVVADEVKALAEESRQSTVVEELVQMAETETVQDLDEVNOR 400

Dh 345 IEAARAGNADGGRSVVADEVKDLAEETODRAMIAAVVEYVATQEDYVTSIQOTRR 404

Qy 401 IGEVVEREAMETLOETITDAVEAASGMOEVSTATDEO-----AVSTEVEAMVD-- 451

Dh 405 VESGSEYESTLRDRIATLADIAEVSNSIDEIORTTSQOAEVTSQATSVSRVAGLSDDT 464

Qy 452 -----GVDRAGEIAALDDIADATDQVFTVEEVEETVGLS 489

Dh 465 TALASDAESAVIGRESAEIATSLQEPONRAVELOQSRVASFVATDESSETAG 518

RESULT 14

HTR2_HALSA STANDARD: PRT: 764 AA.

AC P71410: 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

```

DE Sensory rhodopsin II transducer (HTR-II) (Methyl-accepting phototaxis
DE protein II) (MPP-II).
GN HTR2 OR HTR1I.
OS Halobacterium salinarum.
OC Archaea: Euryarchaeota: Halobacteria: Halobacteriales:
OC Halobacteriaceae: Halobacterium.
OX NCBI_TaxID=2242;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Flx15;
RA MEDLINE=96323203; PubMed=8710852;
RT Zhang W., Brooun A., Mueller M.M., Alam M.;
RT "The primary structures of the Archaeon Halobacterium salinarum blue
RT light receptor sensory rhodopsin II and its transducer, a methyl-
RT accepting protein."
RL Proc. Natl. Acad. Sci. U.S.A. 93:8230-8235(1996).
CC -1- FUNCTION: TRANSDUCES SIGNALS FROM THE PHOTOTAXIS RECEPTOR
CC SENSORY RHODOPSIN II (SR-II) TO THE FLAGELLAR MOTOR. RESPONDS TO
CC LIGHT CHANGES THROUGH THE VARIATION OF THE LEVEL OF METHYLATION.
CC -1- SIMILARITY: BELONGS TO THE CHEMOTAXIS SENSORY TRANSDUCERS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U62676; AAC4369.1; -
DR HSSP: P02942; 1007;
DR InterPro: IPR004089; Cnmtaxis_transd.
DR InterPro: IPR003660; HAMP.
DR Pfam: PF00015; MCPsignal; 1.
DR Pfam: PF00672; HAMP; 2.
DR SMART: SM00304; HAMP; 2.
DR SMART: SM00283; MA; 1.
DR Transducer; Photoreceptor; Transmembrane; Methylation.
FT INIT_MET 0 0
FT DOMAIN 1 15 BY SIMILARITY.
FT TRANSMEM 16 36 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 37 272 POTENTIAL.
FT TRANSMEM 278 298 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 299 764 POTENTIAL.
FT DOMAIN 764 764 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 764 AA; 79187 MW; 1E0D7B4E60FC588 CRC64;

Query Match 19.3%; Score 461; DB 1; Length 764;
Best Local Similarity 28.7%; Pred. No. 2e-13;
Matches 135; Conservative 75; Mismatches 219; Indels 42; Gaps 9;

OY 37 LSEFTGIDDDTMALAAEQLFEATADALVTFYDHLSEYERQDLEFRANSTKIVEQLEKQ 96
DB 294 LSLIGTVGSTYTALRQ--FSRRADEMAAGDDTDIDISR-NDEGTLAESRSMRSL 350
OY 97 AEVLGLGGEYTEYAARARIGKIHVGLGCPDYLYLCATRYTYGGLDALADVYADR 156
DB 351 SEELT-----DAERATARAEDAR-EDAEQORADA-----EAAREDEAAR 389
OY 157 --GEAAAVDELVARFELPKLTLT--FDQOIAMDTYIDSYAQRLL-----DEIDSQ 205
DB 390 KDAQETARLLESAAADYEALTRVADGDLTRRYDASRDHDAKRIHALNDMLDILETSV 449
OY 206 ELANAVATHEAEPPLSLLEA-----TSQDVARTDTMRARDDQVDRMADVSREISSV 258
DB 450 AAATARSADVSDAAQGVEDADAGDAIDAGTDVSTAVDEISDGAETGDRLEHVAAGEVDLS 509
OY 259 ASVEEVASTADYRTSEDAEALAOGEAAADALATMTDIDEAIDVTVAGVEQIGERRA 318
DB 510 ASAEVEAEVTSADTACGAASAVDDGROATEDAVETMDVDADDAADAMALDSEMA 569
OY 319 DVESVYGVDDIAEQTNNALNASTAARAGEGEGFAVVADEVKRLAESRQSTRVEE 378
DB 570 DIGEIVDVIADIDQTNMLNASTAARTGADGDFAVVADEVKTLAESRDAEDIES 629

```

```

OY 379 LYBQQAETEETVDQDDEVNQRIGEVEREAMETLQETIDVADPAASGMQEVSTATDE 438
DB 630 RLALAGQSDVADDEKRAVSDTVSGRAVFGAATFLDDVSVSVADTDAAGIRATDR 669
OY 439 QAVSTEVAEMVDGVDRAGEIAALDDIADATDQGVRYVEEYRETVKLS 489
DB 690 QAHAASRVASAVDEAVAGISQETAAQATAVADSAAATQDTLSSVDMAADLA 740

RESULT 15
HTR1_HALSA
ID HTR1_HALSA STANDARD; PRT; 535 AA.
AC P33955;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sensory rhodopsin I transducer (HTR-I) (Methyl-accepting phototaxis
DE protein I) (MPP-I).
GN HTR1 OR HTR1 OR HTR.
OS Halobacterium salinarum.
OC Archaea: Euryarchaeota: Halobacteria: Halobacteriales:
OC Halobacteriaceae: Halobacterium.
OX NCBI_TaxID=2242;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=L33;
RA MEDLINE=90384855; PubMed=2205842;
RT Ferrando E., Krah M., Martan W., Oesterhelt D.;
RT "A gene from S. pombe with homology to E. coli RNase III blocks
RT conjugation and sporulation when overexpressed in wild type cells."
RT Nucleic Acids Res. 18:5304-5304(1990).
RN (2)
RP FUNCTION.
RX MEDLINE=94244615; PubMed=8187768;
RA Krah M., Martan W., Verneiglio A., Oesterhelt D.;
RT Phototaxis of Halobacterium salinarum requires a signalling complex
RT of sensory rhodopsin I and its methyl-accepting transducer HtrI."
RL EMBO J. 13:2150-2155(1994).
CC -1- FUNCTION: TRANSDUCES SIGNALS FROM THE PHOTOTAXIS RECEPTOR
CC SENSORY RHODOPSIN I (SR-I) TO THE FLAGELLAR MOTOR. RESPONDS TO
CC LIGHT CHANGES THROUGH THE VARIATION OF THE LEVEL OF METHYLATION.
CC -1- SIMILARITY: BELONGS TO THE CHEMOTAXIS SENSORY TRANSDUCERS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X68591; CAA48578.1; -
DR PIR: S28466; S28466.
DR HSSP: P02942; 1007;
DR InterPro: IPR004089; Cnmtaxis_transd.
DR InterPro: IPR003660; HAMP.
DR Pfam: PF00015; MCPsignal; 1.
DR Pfam: PF00672; HAMP; 1.
DR SMART: SM00304; HAMP; 2.
DR SMART: SM00283; MA; 1.
DR Transducer; Photoreceptor; Transmembrane; Methylation.
FT INIT_MET 0 0
FT DOMAIN 1 8 BY SIMILARITY.
FT TRANSMEM 9 29 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 30 33 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 34 54 POTENTIAL.
FT TRANSMEM 55 54 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 55 535 CYTOPLASMIC (POTENTIAL).
FT MOD_RES 265 265 METHYLATION (BY SIMILARITY).
FT MOD_RES 272 272 METHYLATION (BY SIMILARITY).
FT MOD_RES 279 279 METHYLATION (BY SIMILARITY).
FT MOD_RES 463 463 METHYLATION (BY SIMILARITY).
FT MOD_RES 472 472 METHYLATION (BY SIMILARITY).

```

SO SEQUENCE 535 AA; 56814 MW; 504A165A47FA8A45 CRC64;

Query Match 18.8% Score 449.5; DB 1; Length 535;
Best Local Similarity 27.7%; Pred. No. 4.4e-13;
Matches 141; Conservative 86; Mismatches 191; Indels 91; Gaps 15;

```
QY 28 LDEAIAWRSLFTGIDDTMAALAEQPLFEATADALVTFDHLSEYERTODLFANSTK 87
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 51 INAGETGRIGKIGAGTGRVANGNLEQEVSTRID---EFGSLADSIQMRQSLRGRLN 106
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 88 TVEQ---LKTQAEYLGLGRGEYDTEYAQAQRARIGIHVYLGIPVYLGAVTRYTG 143
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 107 EMERTSRDLEETQTE---SETTRESEQAQKQASAE-----REARE 145
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 144 LLDALADVADRGE--PAAAVDELVARFLPMLKLTFFDOI-----AMDTYIDSYA 194
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 146 LAATYQDSSNRD-GETMESSATGD-----FTQRVDVDTDPKEMETFGTA FN 190
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 195 QRLHDEIDSROELANAVATHEAPLSLEATSODV-AERTDTMRA-----RTDDQVDRM 247
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 191 QMM-DLQATVRYTVTVADEIEAKTERMSSETADIEASAGDTEAVSKIESQPNDRTEL 249
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 248 ADVGREISSVASVEEVASTADVARTSEDAEALAQGEAAADALATMTDIDEATDGV 307
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 250 DSADDVQVQVSAASEIAATIDDLASRESDVATASDAARDSSKSLDEMSSIETEVDPV 309
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 308 AGVQLGRADVDSYGVVIDIAEQTMALNLSIEARAG--EAGEFAVVADEVKAL 365
   ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 310 GQVHLRDQVAEITDIVIYIDIGEQTMALNLSIEAARAGNADGDFSVVADEVKDL 369
   ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 366 AEESSREOSTREVELEVEQMAETETVDQLDDEVNORIGEVERVEEAMETLOETDAVEDA 425
   ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 370 AEFODRANETIAAAYEKVTAQIEDVTASIOHTRRVESGSETVESTLHIRTADSISEV 429
   ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 426 ASGMEVSTATDEQ-----AVSTEVEAEMVD-----GVDDRAGEIAAAL 464
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 430 SNSIDEIORTTSEQAETVQSTATSEYRVAAGLSDDTTALASDAESPVIGORESAEEIAASL 489
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 465 DDIDATDQOVRT-----VEEVRETVG 486
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 490 EQFONTAVEQLQSRVASTVATEDSETAG 518
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

Search completed: January 2, 2003, 12:41:59
Job time : 16 secs

1	359.5	15.0	891	10	US-09-272-0630-C	Sequence 5, Appl1
2	292.5	12.2	531	1	US-08-976-0639-C-34	Sequence 34, Appl
3	254	10.6	433	8	US-08-945-038-6	Sequence 60, Appl1
4	239.5	10.0	2310	9	US-09-991-446-120	Sequence 120, App
5	239.5	10.0	2310	10	US-09-874-923-120	Sequence 120, App
6	194	8.1	1786	9	US-09-742-056-3	Sequence 3, Appl1
7	190	7.9	1879	9	US-09-971-536-70	Sequence 70, Appl1
8	187	7.8	630	9	US-09-742-096-5	Sequence 5, Appl1
9	177	7.4	1179	10	US-09-815-242-13608	Sequence 13608, <i>n</i>
10	176	7.4	540	9	US-09-742-096-32	Sequence 22, Appl1
11	175.5	7.3	1289	9	US-09-712-363-359	Sequence 259, App
12	175.5	7.3	1798	10	US-09-845-583-8	Sequence 8, Appl1
13	174	7.3	1179	10	US-09-815-242-13262	Sequence 13262, <i>n</i>
14	173.5	7.2	1798	10	US-09-938-275-9	Sequence 9, Appl1
15	173	7.2	1786	10	US-09-873-676-113	Sequence 113, Appl1
16	173	7.2	1786	10	US-09-938-275-6	Sequence 6, Appl1
17	171.5	7.2	537	10	US-09-815-242-13463	Sequence 13463, <i>n</i>
18	171.5	7.2	1786	10	US-09-938-275-7	Sequence 7, Appl1
19	171	7.1	2478	10	US-09-815-242-5816	Sequence 5816, <i>n</i>

20	171	7.1	2478	10	US-09-815-242-12967	Sequence 12967, AP
21	169.5	7.1	1841	10	US-09-825-299-911	Sequence 911, AP
22	169	7.1	1202	10	US-09-838-273-8	Sequence 8, AP
23	168	7.0	1607	10	US-09-838-275-10	Sequence 10, AP
24	167.5	7.0	2368	10	US-09-815-242-5635	Sequence 5635, AP
25	167.5	7.0	2368	10	US-09-815-242-12389	Sequence 12389, AP
26	167.5	6.9	1609	10	US-09-838-275-11	Sequence 11, AP
27	163.5	6.8	1799	10	US-09-845-583-6	Sequence 6, AP
28	163.5	6.8	2076	10	US-09-815-242-5815	Sequence 5815, AP
29	163.5	6.8	2186	10	US-09-815-242-12913	Sequence 12913, AP
30	161	6.7	1155	9	US-09-738-526-5764	Sequence 5764, AP
31	159.5	6.7	600	9	US-09-738-526-5167	Sequence 5167, AP
32	159	6.6	1925	12	US-10-028-946-4	Sequence 4, AP
33	159	6.6	2058	10	US-09-815-242-5703	Sequence 5703, AP
34	159	6.6	2053	9	US-10-017-916-7	Sequence 7, AP
35	159	6.6	2054	12	US-10-028-946-2	Sequence 2, AP
36	159	6.6	3158	10	US-09-815-242-12611	Sequence 12611, AP
37	158	6.6	11962	10	US-09-815-242-11828	Sequence 11828, AP
38	154	6.4	1597	9	US-10-017-916-6	Sequence 6, AP
39	154	6.4	2055	9	US-10-017-916-4	Sequence 4, AP
40	153.5	6.4	2437	10	US-09-815-242-5834	Sequence 5834, AP
41	153.5	6.4	6281	10	US-09-815-242-12996	Sequence 12996, AP
42	153	6.4	2835	10	US-09-885-535-4	Sequence 4, AP
43	152	6.3	396	10	US-09-800-729-207	Sequence 207, AP
44	150.5	6.3	463	9	US-09-738-626-9970	Sequence 3970, AP
45	150.5	6.3	566	9	US-09-991-496-8	Sequence 8, AP

ALIGNMENTS

RESULT 1
MS-00-073-000-E

Patent No. US20020022239A1

; GENERAL INFORMATION:

APPLICANT: Lagarias, John C.

; TITLE OF INVENTION: Phytofluors as fluorescent labels

FILE REFERENCE: 2500.118US0

CURRENT APPLICATION NUMBER: US/

CURRENT FILING DATE: 1998-03-19

NUMBER OF SEQ ID NOS: 24

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 5

LENGTH: 891

TYPE: PRI
ORGANISM: Malesona

ORGANISM: ULIKIOWITZ

OTHER INFORMATION:

OTHER INFORMATION

OTHER INFORMATION

OTHER INFORMATION

US-09-272-809-5

0
1
2
3
4
5
6
7
8
9

query match

Best Local Similarity

Matches 136; Cons

QY 23 ADRIGLDEAET

Db 4.1.8 ADRV-----I

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

QY 76 ERTQDLFANSI

Computer Mac

DD 470 QUALITY - NAG

126 -----CD-----

227 677 GP 11

513 АНОНСОВИ

DO NOT REPRODUCE OR TRANSMIT

OV 174 MIKILTFDOOT

[illegible]

571 L--LMEVDPVS

QY 234 DT-----MRATDQVDMADYSREISSVSASVEEVASTADVVRTSDEBALAQGE 286
DB 627 DTNEVAVROLAQOANRQALDVALEBRLQMNKSIQAVNAQAESAQVRAQTVQDGE 686
QY 287 AADDAALATMTDIDEATDGTAGVEOLGERAAVESTGVYIDIAEOTNNLALNASTEAA 346
DB 687 DAAKRTVDGIVARETVAAATAKQVKRGESSQKISKVNNLIGSFAQOTNLALNAAIEAA 746
QY 347 RAGEAGEFAVVADEVKALAEESREOSTREVEELVEOMQAEETEVOLDDEVNORIGGYE 406
DB 747 HAGEERGFVVADEVRSALARQSAEATAEIAQVLVATIGATNEVNAEMAGTEQVVGTR 806
QY 407 RVEAMETLOETIDAVEDAAGMOE-VSTATDEQAVSTEE-----VAEMVD----- 451
DB 807 LIVEETRSLNQIT-AVSAQISGLVEAITSMAIEQSOTSESQTMALVQAIDKNSEAS 865
QY 452 GVDPRAGEIAALDIDADTDQ 473
DB 866 GVSATFKELAAVQSLQEAQV 887

RESULT 2

US-08-976-063C-34
Sequence 34, Application US/08976063C
Publication No. US20020182657A1
GENERAL INFORMATION:
APPLICANT: Alexander Steinhuchel; Horst Priefert; Jurgen Rabenhorst
TITLE OF INVENTION: SYNTHETIC ENZYMES FOR THE PRODUCTION OF
TITLE OF INVENTION: CONFERTYL ALCOHOL, CONFERTYLALDEHYDE, FERULIC ACID, VANILLIN A
NUMBER OF INVENTIONS: 42
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SPRUNG KRAMER SCHAEFER & BRISCOE
STREET: 660 White Plains Road
CITY: Tarrytown
STATE: New York
COUNTRY: U.S.A.
ZIP: 10591-5144
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB storage
COMPUTER: HP VECTOR
OPERATING SYSTEM: DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/976, 063C
FILING DATE: 21-NOV-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 196 49 655.1 (Germany)
FILING DATE: 29-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kurt G. Briscoe
REGISTRATION NUMBER: 33,141
REFERENCE/DOCKET NUMBER: Beyer 9998-CAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (914) 332-1700
TELEFAX: (914) 332-1844
TELEX:
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-976-063C-34

Query Match 12.2%; Score 292.5; DB 1; Length 531;
Best Local Similarity 24.7%; Pred. No. 1.1e-10;
Matches 110; Conservative 86; Mismatches 192; Indels 57; Gaps 11;
QY 63 ALVTDFYHLESEYERRQDL-FANSTKYVEQJKEHQAEVLLGLGKEVTEVTEAAQARATG 120

DB 90 AEVNDALIDFDFQGRSELLKFAASSR-----EASYSIEVG-----SMAADKLQSG 136
QY 121 KIHVDVLGPGDVLGATRYRTYGLDLADLVADRGEE-----AAA 162
DB 137 EPSDAL-MVADKRLNVEYEDLSSAVNRLMGLHLEDONEKPLIYMLGVTTLTMLMSAY 195
QY 163 ADELVAARFLPMUK-LTFPDQJAMDYIIDSVAORLHLEDISQOELANAVATVHEAPLSS 221
DB 196 SVMFISRQVLPPLKSTVQLAERIAISGDLADVGDSSRRDEIQLOQSATRRMAIGLRNLVGD 255
QY 222 L-EATSDQVAERTDTRARTRDQVD-----RAADVSRREISSVSASVEEVASTADVVRT 274
DB 256 IGSRAQVLSSSDLSAICAQAQIDVECOKLSVAQVSTAVNELVETVQALAKSTEEAATV 315
QY 275 SEDAEALAQGEAADALATMTDIDEATDGTVA-----GVEQLGERAADESVTGVIDI 330
DB 316 AVLADEKARGESVYKAV-----DTEHLSGDMAELGDAMERQONSAQINKVYDIKAV 371
QY 331 AEOTNNLALNASTEAARAGEGFAVVADEVKALAEESREOSTREVEELVEOMQAEET 390
DB 372 AEOTNNLALNASTEAARAGEGFAVVADEVKALAEESREOSTREVEELVEOMQAEET 431
QY 391 VDQDEVNORIGGVERVEAMETLOETIDAVEDAAGMOEVSSTATDEQAVSTE----- 444
DB 432 GELMRGKVRTHDVGLAQQAARATRNYPVAVGIAAMVQIAAGAEQGAAYVQIIONM 491
QY 445 -EVAEMVDGVDPRAGEIAAALDIDIA 468
DB 492 LEVHKMADESATKAQOTMKSKEIA 516

RESULT 3

US-08-945-038-6
Sequence 6, Application US/08945038
Patent No. US20020146423A1
GENERAL INFORMATION:
APPLICANT: Doidge, Christopher V.
APPLICANT: Lee, Adrian
APPLICANT: Radcliff, Fiona J.
APPLICANT: Hocking, Dianna M.
APPLICANT: Webb, Elizabeth A.
TITLE OF INVENTION: PROTECTIVE HELICOBACTER ANTIGENS
NUMBER OF INVENTIONS: 26
CORRESPONDENCE ADDRESSES:
ADDRESSEE: FOLEY & LARDNER
STREET: 3000 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/945, 038
FILING DATE: 23-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU96/00225
FILING DATE: 19-APR-1996
APPLICATION DATA:
APPLICATION NUMBER: AU PN2575
FILING DATE: 21-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PN3931
FILING DATE: 03-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PN7565
FILING DATE: 16-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 017227/0133
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 433 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-945-038-6

Query Match 10.0%; Score 254; DB 8; Length 433;
Best Local Similarity 29.3%; Pred. No. 1.7e-08;
Matches 80; Conservative 48; Mismatches 97; Indels 48; Gaps 8;

QY 191 DSYAQRLEHDEISROELANAVATHVE-----APLSLEATS---ODVAE 231
DB 15 DSEIAELKKEVNLKYOSLLN-LCLHEGFVGIKNNKVFYKSGNLASLNLEEQSVHFKENAE 73
QY 232 RFD-----TMRARDDOVDRMADVSRE-----ISSVSASVEEASTADVRR 274
DB 74 SYDLQGVSYLSKSQNDIGVQYFSLAKKTCVGEYHKNDLFKTFCASLKEGLENAQESMXY 133
QY 275 SEDAEAL-----AOGGEAADDALATWT---DID---EATDGVTAGVEOLGERAAVESV 323
DB 134 FHQETGLLNAAKNGEASHTEGLGTYNKTCGQDIESLEYEKQNATSLADSLNRSNETIYV 193
QY 324 TGVVIDIAEQTMALNLSIEARAGEGEGFAVVADEVKALAEESREOSTRVEELVEOM 383
DB 194 ISLIDIDIAEQTMALNLSIEARAGEHGRGFAVVADEVKRLAEKTKETKEIVVVKSM 253
QY 384 QAEETFTVDQDEVNORIGEGYERVEAMETLO 416
DB 254 QOEANDIQTNTHDINSIVSSIKGDVEELKSTYK 286

RESULT 4
US-09-991-496-120

; Sequence 120, Application US/09991496
; Patent No. US20020169285A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, David C.
; APPLICANT: Skelky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Colier, Rhea
; APPLICANT: Probst, Peter
; APPLICANT: Brannon, Mark
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS
; FILE REFERENCE: 210121.420C9
; CURRENT APPLICATION NUMBER: US/09/991,496
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 120
; LENGTH: 2310
; TYPE: PRT
; ORGANISM: Leishmania major and chagasi
US-09-991-496-120

Query Match 10.0%; Score 239.5; DB 9; Length 2310;
Best Local Similarity 22.9%; Pred. No. 8.1e-07;
Matches 121; Conservative 86; Mismatches 224; Indels 97; Gaps 18;

QY 14 RNCIDSHALADRIGLDEAETAMRLSTGIDDTMAALAEQPLFEATADALVDFYDHE 73
DB 960 RLAAADDEARQOLAAVAEEIQRL-----DTATQORAE---LEAQLARLAAD----- 1003
QY 74 SYERTQDDEANSTKTYEQAKETQAEYLLGLGREGYDTEYAAQARARKIGHVGLGRPVY 133

DB 1004 RDEARQOLAAAEELQRLDTATQO-----RAELEAQVRLAAAEELQRLDTATQO-- 1055
QY 134 LGATYTRYTGLLDALADVVADRG---EAAANVDELVARFLPMKLTFFDOOIAMDYI 190
DB 1056 -----QRAELEAVARLAADDEARQOLAAAEELQRL-----LDATQORAE---EL 1099
QY 191 DSYAQRLEHDEISROELANAVATHVEAPLSLEATSODVAERTDT---MRARDDOV 244
DB 1100 EARVARLAANAEELQRLDTATQORAELEAVARLAANAEELQRLDTATQORAELEAV 1159
QY 245 DRMA---DVSREISSVSASVEEVASTAD-----DVRTSDEBALAQ-----Q 284
DB 1160 ARLAADDEARQO---QLAANAEELQRLDTATQORAELEAVARLAANAEELQRLDTATQ 1217
QY 285 GEAAADALATM-TDIDEATDGVTAGVEOLGE-----AADVESVG 325
DB 1218 QRAELEAQVRLAAADDEARQOLAAAEELQRLDTATQORAELEAVARLAANAEELQ 1277
QY 326 VIDIDIAEQTMALNLSIEARAGEGEGFAVVADEVKALAEESREOSTRVEELVEOM 385
DB 1278 RUDTATQORAELEAVARLAADDEARQOLAAAEELQRLDTATQORAELEAVARLA 1337
QY 386 ETEETVDQL-----DEVNORIGEGYERVEAMETLOEITDAVEDAASGMEVSTATDEGAV 441
DB 1338 DQDEARQOLAAAEELQRLDTA---TQORAELEAVARLAANAEELQRLDTATQORAE 1394
QY 442 STEEVAEMVDGDDRAGEITAAALDIDATDQOVRFVEEYRETVGKLS 489
DB 1395 LEARVARLAADDEARQOLAAAEELQRLDTATQORAELEAVARLA 1442

RESULT 5
US-09-874-923-120

; Sequence 120, Application US/09874923
; Patent No. US20020081320A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, David C.
; APPLICANT: Skelky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Colier, Rhea
; APPLICANT: Probst, Peter
; APPLICANT: Brannon, Mark
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS
; FILE REFERENCE: 210121.420C8
; CURRENT APPLICATION NUMBER: US/09/874,923
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 120
; LENGTH: 2310
; TYPE: PRT
; ORGANISM: Leishmania major and chagasi
US-09-874-923-120

Query Match 10.0%; Score 239.5; DB 10; Length 2310;
Best Local Similarity 22.9%; Pred. No. 8.1e-07;
Matches 121; Conservative 86; Mismatches 224; Indels 97; Gaps 18;

QY 14 RNCIDSHALADRIGLDEAETAMRLSTGIDDTMAALAEQPLFEATADALVDFYDHE 73
DB 960 RLAAADDEARQOLAAVAEEIQRL-----DTATQORAE---LEAQLARLAAD----- 1003
QY 74 SYERTQDDEANSTKTYEQAKETQAEYLLGLGREGYDTEYAAQARARKIGHVGLGRPVY 133
DB 1004 RDEARQOLAAAEELQRLDTATQO-----RAELEAVARLAANAEELQRLDTATQO-- 1055
QY 134 LGATYTRYTGLLDALADVVADRG---EAAANVDELVARFLPMKLTFFDOOIAMDYI 190

Db 1056 -----QRAELARVARLADREARQOLAAAEELQOR-----LDTATQORA---EL 1099
Qy 191 DSYAQR-----HDEIDSRQELANAVATHEAPLSLEATSDVABRTD---MRARTDOV 244
Db 1100 EARVARLAAAEELQORLDTATQORAELEAOVARLAAAEELQORLDTATQORAELEARV 1159
Qy 245 DRMA---DVSREISSVASVEEASTAD-----DVRTSEDAEALAQ-----Q 284
Db 1160 ARLAADREARQ---QLAAAEELQORLDTATQORAELEAOVARLAAAEELQORLDTATQ 1217
Qy 285 GEAAADALATM-TDIDEADTGVATGVEQGER-----ADVESYTG 325
Db 1218 QRAELQOLALADREARQOLAAAEELQORLDTATQORAELEAOVARLAAAEELQ 1277
Qy 326 VIIDIAQOTNLALNASTEARAGEGFAVVADEVKALAEESREOSTRYEEVEOMQA 385
Db 1278 RLDTATQORAELEARVARLADREARQOLAAAEELQORLDTATQORAELEARVARIAA 1337
Qy 386 ETEETVDQ---DEVNORIGEGVERVEEAMETLQETITDAVEDAASQOEVSATDQAV 441
Db 1338 DGDEARQOLAAAEELQORLDTA---TQORAELEAOVARLAAAEELQORLDTATQORAE 1394
Qy 442 STEEVAEMVQYDDRAGEIAAALDDIDADIDQVRYVEEARETYGKS 489
Db 1395 LEARVARLADREARQOLAAAEELQORLDTATQORAELEAOVARLA 1442

RESULT 6

US-09-742-096-3
; Sequence 3, Application US/09742096
; Patent No. US20020135441A1
; GENERAL INFORMATION:
; APPLICANT: DRUIHE, PIERRE
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-BRYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 2007730SDIY
; CURRENT APPLICATION NUMBER: US/09/742,096
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/973,642
; PRIOR FILING DATE: 1998-02-06
; PRIOR APPLICATION NUMBER: PCT/FR96/00894
; PRIOR FILING DATE: 1996-06-12
; PRIOR APPLICATION NUMBER: FR 95/07007
; PRIOR FILING DATE: 1995-06-13
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 3
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: P. falciparum
US-09-742-096-3

Query Match

Best Local Similarity 8.1%; Score 194; DB 9; Length 1786;
Matches 100; Conservative 85; Mismatches 139; Indels 84; Gaps 22;

Qy 146 DALADVY---VADGEEAAA-VDELVARFLPMKLTLTFQOIAMDTYIDSYAQRLEDEI 201
Db 565 ESAVAENVEESAENVEESAENVEIYAPVEIYAPVEIYAPVEIYAPVEIYAPVEIYAPVEIY 623
Qy 202 DSRQELANAVATHEAPLSLEATSDVABRTDQVDRMADVREI-----S 255
Db 624 EENVE---ESAENVEE---SVAENVEESAENVEIYAPVEI---VAPTEEIVAPSVVE 677
Qy 256 SVASVEEVA-----STADVVRT-SEDAEALAQGEAAADALATMTP-IDERATGCVT 307
Db 678 SVAPSVESVEENVEESAENVEESAENVE-----ESAENVEESAENVEI---VA 728
Qy 308 AGVEOLGERAD-----VESYTGVIDTAEQTNMLALNASTEARAGEGFA----- 356
Db 729 PTVEEIVAPVEIYAPVEIYAPSVESVAENVEESAENVEESAENVEESAENVEESAENVE 788
Qy 357 -VVADEVKALAEESREOST--RYEEVEOMQA-----ETEETVD---QL 394

Db 789 ESVAPTVEEIVAPSVESVAPSVESVAENVATMLSDNLSNLGJETEIEKSIINEI 848
Qy 395 DEVNORIGEGV-EVEEVA-----METLOEIT---DAVDAASQME-VSTATDE 438
Db 849 EEVENEVNTILLEVEETFAESVTFSTFNILEIDENTITNDTIEKEEHEENVLSALE 908
Qy 439 QAVSTEVAEMVQYDDRAGEIAAAL-DDIADATDQVRYVEEARETY 485
Db 909 NTQSEEREKVEYIDYIEEVEKAVATTLTETVQAEKSNATITTEFENL 956

RESULT 7

US-09-971-536-70
; Sequence 70, Application US/0971536
; Patent No. US20020159976A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Ilkka
; APPLICANT: Bloksberg, Leonard
; APPLICANT: Lubbers, Mark
; APPLICANT: Dekker, James
; APPLICANT: Christenson, Anna
; APPLICANT: Holland, Ross
; APPLICANT: O'Toole, Paul
; APPLICANT: Reid, Julian
; APPLICANT: Coolbear, Timothy
; TITLE OF INVENTION: Lactobacillus rhamnosus Polynucleotides, Polypeptides and Meth
; FILE REFERENCE: 1043c2
; CURRENT APPLICATION NUMBER: US/09/971,536
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: U.S. No. US20020159976A1 09/634,238
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: U.S. No. US20020159976A1 09/724,623
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: PCT/NZ01/00160
; PRIOR FILING DATE: 2001-08-08
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 70
; LENGTH: 1879
; TYPE: PRT
; ORGANISM: Lactobacillus rhamnosus
US-09-971-536-70

Query Match

Best Local Similarity 7.9%; Score 190; DB 9; Length 1879;
Matches 119; Conservative 83; Mismatches 185; Indels 168; Gaps 25;

Qy 12 DVKNGIDGHALADRIQDEAEIAWRLSFTGIDDTMAALAEOPLPFATADA----- 63
Db 992 DAKKAITAEK-----AKVT-----KAIDOPTILTAEEKKQKQAVDAETKAKAAI 1037
Qy 64 ---LVDPFYHLESYERTODLFANSTKYTVQOLKETOAEYLLGIGRGEYDTEYAQR--- 116
Db 1038 DATLVADAID-----QALADGKITID-----AQYOTGIA---LDRKKAAKOTID 1079
Qy 117 ARIGKIDVGLGPDVYLGATRYTGLDALADVDVADGEAAVAADVELVARFLMLK 176
Db 1080 ABAKASE---AIDDDVLTLDQK-----ATQKQAAVDAETKAKAIDQ----- 1120
Qy 177 LTFDQOIAMDTYIDSYAQRLEHDEIDSRQELANAVATHEAPLSLEATSDVABRTD-- 234
Db 1121 --ASDANAVQATIDG-----EAIHQHSATRLDQKQKQAKAIDAEBAKVSKALDQD 1173
Qy 235 -----TWARTDQVDRMADVSRE-----ISSVASVEEVASTADV 271
Db 1174 VTLTATQKADQKQAVIADKAKKLIDAGNADGIKQAESDGIKAIDAQHQSSQALADRK 1233
Qy 272 R--RTSEDAAL-----AQGGAADDLALATMTDDEA----- 302
Db 1234 RDKATTAIDAEARETAIDHDATLTANEKASQKQAVTDEATKAKKAKAIDAKQADAVDQAK 1293

QY 303 TDGVTA-----GVEQLGERADVESVTGVID-DIAEOTNMLALNASTIARAGEAGEGF 355
DB 1294 TGIKAIQIDQHHSGALDRKADAKO---VIDAEAKYTAALDQDNTLTKAKQAOKQGV 1350
QY 356 AVADEVKALAEESREQSTRVEELVEQMOA-----ETEEVTVD-QLDEVNQRIGEGERVE 409
DB 1351 ATEADKAKQAIDAGADADAVDQKTAGIQIDAIDQHKAGTIDSRHDDAKQAIDEAKYI 1410
QY 410 EMETLQETTDVADEDAASQMEVSTATDQAVSTEVAMVDGVDRAEIIAALDDIAD 469
DB 1411 KAIIDDPITLTAQOKA-----QKQAVATE--ADKAKKAIADAGD-----AD 1449
QY 470 ATDQO-----VTRVEE 480
DB 1450 AVDQAKTAGIKAIKAI 1464

RESULT 8
US-09-742-096-5
; Sequence 5, Application US/09742096
; Patent No. US20020155441A1
; GENERAL INFORMATION:
; APPLICANT: DROULHE, PIERRE
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 200773USODIV
; CURRENT APPLICATION NUMBER: US/09/742,096
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/973,642
; PRIOR FILING DATE: 1998-02-06
; PRIOR APPLICATION NUMBER: PCT/FR96/00894
; PRIOR FILING DATE: 1996-06-12
; PRIOR APPLICATION NUMBER: FR 95/07007
; PRIOR FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 630
; TYPE: PRT
; ORGANISM: P. falciparum
US-09-742-096-5

Query Match 7.8%; Score 187; DB 9; Length 630;
Best Local Similarity 21.5%; Pred. No. 0.00025;
Matches 85; Conservative 87; Mismatches 138; Indels 86; Gaps 16;

QY 150 DDVADRCGEAAA-AVDELVARFLPMLK--LLTFDQIAMDYIDSYAORLHDEIDSROE 206
DB 263 EEIVAPTVESVAPYEEIYA---PIVEESVAPYVEIYPPSVESVAVESV--AEN 317
QY 207 LANAATVHEAPLSLEATSOVAERTDTMRARTDQDVRMADVSRHSSVSASVEVA- 265
DB 318 VEESVAENEE--SVANEEVESAENVEESVAENVEI-----VAPSVETIYA 363
QY 266 -----STADVARTSDEADALAO-----QGEAADAADLTMDIDATGCVAGVQOLE 315
DB 364 PVEESVAVANVATNLSL--NLSNLLGIGITEIKISILNEIEVKF--NVVTTLKEVBE 420
QY 316 RAAD--VESYTGVIDIAEOT-----NMLANASTIARAGEAGEGFVAVADE 361
DB 421 TTAESTYTSNLTLEEIQNTIINDTIEKLELHENVLTAALENQSEKEKVIDIIEE 480
QY 362 VVALAEESREOSTRVEELVEQMOAETEEVDOL-----DEVNQRIGEGERVEAM- 412
DB 481 VR-----EVAATTLTIEVEQAEESSESYITTEIFENLSENAVESNEKVAENLEKLTETVF 534
QY 413 -ETLQETITAVE-----DAASQMEVSTATDQAVSTEVAMVDGVDRAEIIAALDDIAD 451
DB 535 NVVLQKVEETVELSGESLENEMDKAFSEIFDNVAKGIOENLLTGMRISLETISYIQSEB 594
QY 452 GVDNRAGEIAAALDDIADATDQAVRTVEEVRGVK 487

DB 595 KVDLNNVVSILDNIEENKKEGLNKLNNISSTEGE 630

RESULT 9
US-09-815-242-13608
; Sequence 13608, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes In
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13608
; LENGTH: 1179
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13608

Query Match 7.4%; Score 177; DB 10; Length 1179;
Best Local Similarity 18.8%; Pred. No. 0.002;
Matches 126; Conservative 114; Mismatches 206; Indels 224; Gaps 28;

QY 30 EAEIARLSTFGTIDDDTMAALAAOPLFEATBAVLTDFYDHLIS--YERTODLFANSTKT 88
DB 364 EAEI--LAFSD--DPQOMTELLRERFVALLQERADVSNQUTRIENLENSRQLSOKQADO 419
QY 89 VROLKE-----TQAEYL-----LGLRGEYDTEYAAO----- 115
DB 420 LERLKRQALIAKEKASQOELEETAKEQYOKLADYQATAKBEQKTSYQAQOQLFDR 479
QY 116 -----RARIKIHVLDLGPVYIGAT-----RYTGL- 144
DB 480 LNLKKNQARAQSLNLTNNHNSFVAGVKSVLQEKDRLGIGTGAVSEHLTFDMHQTAL 539
QY 145 --LDALADVVADRCGEAAAADVELV-----ARFLPM--LKLTP----- 180
DB 540 IALGASSCHIIYEDENAAKRAIDFLKRNKRAFTPLPITTIKARTISSONQDAIIVSPGF 599
QY 181 ----DQOIAMDYIDSYAORL-----HDEIDSROELANAVATVHEAPLSLEATSQDVA 230
DB 600 LGMADLVTFDRLEAIFKRLNLATTAIFPTVEAREAAQVRYOVR--NVTLTGTEL-- 654
QY 231 ERTDITRA-----RTDQV--DRMADVSRHSSVSASV--EVA--STADVARTSDEAE 279
DB 655 -RTGGSYAGCANRQNNSTIKPELEQLOKEIAEESLSLSEVALTKTLQDEMARLTSELE 713
QY 280 ALAQOGEAA-----ADDALATTTTDI--DEATDGVTA-----VEQL 313

Db 714 AITSQSGQARIQEGGLEFLAYQOTSQOYVELETLMKLOEEIDRLSEGDQWQAKENCQERL 773
QY 314 GERADAEVSTGVDDIAEOTNMLA---LNASTEPARA-----G 349
Db 774 AALASQONLEAELEEKSKKNKAIQERYQNLQOEELQARLLKTELQCKREYADIERLG 833
QY 350 EAGEGFAVNADEVKAL-----AESSEOSTREVELVEQWQAEETEET 390
Db 834 KEIDNLNIEOEIORMLOEKVNDLEKVDLNLQQAEEAKTKQTNLQOGLIRKQFELDDI 893
QY 391 VDQUDENQNRIGES-----VERVEAMETLQ-EITDANEDAASQOEY 432
Db 894 EGOIDDASHLDQARQONEEMIRKQTRAEAKKEKVESERLHLSQLDQYQISTEALBK 953
QY 433 STADEQAVSTEVEAEM-----VDGVD-----DRAGEIAALDDIADATDOQV 475
Db 954 AHELENINLAEQEKDLEKAIKRSIGPVNINAIQOYEVEYKRLDPLNSQRDILSAKMLL 1013
QY 476 RTVEEYRETV 485
Db 1014 ETTEMNDEV 1023

RESULT 10
US-09-742-096-22
; Sequence 22, Application US/09742096
; Patent No. US20020155441A1
; GENERAL INFORMATION:
; APPLICANT: DRUIHE, PIERRE
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 200773050D1V
; CURRENT APPLICATION NUMBER: US/09/742,096
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/973,642
; PRIOR FILING DATE: 1998-02-06
; PRIOR APPLICATION NUMBER: PCT/FR96/00894
; PRIOR FILING DATE: 1996-06-12
; PRIOR APPLICATION NUMBER: FR 95/07007
; PRIOR FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 540
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-09-742-096-22

Query Match 7.4%; Score 176; DB 9; Length 540;
Best Local Similarity 25.7%; Pred. No. 0.00095;
Matches 96; Conservative 78; Mismatches 142; Indels 58; Gaps 20;

QY 146 DALADY---VADRGEEAANA-VDELVARFLPMKLLTFPOQIAMDYIDSVQRLHDEI 201
Db 39 ESVAENVEESVAENVEESVAENVEEIVAPTVEIAPTVEIIVA-PSVESVAPSVESV 97
QY 202 DSNQELANAVATHVEAPL-----SILEATSOQVAERTDTMRARTDOQVRMADVSREIS 255
Db 98 EENVE--ESVAENVEESVAENVEESVAENVEESVAENVEEIVAPTVEI 152
QY 256 SVASAVEE-----VASTADVVRTSED--AEALAOQGEAA---ADDALATMTDIDEAT 303
Db 153 -VAPTVEIAPSVESVAPSVESVAPSVESVAPSVESVAPSVESVAPSVESVAPSVESV 211
QY 304 DGVTAGVQLGEAAD--VESVTGVDDIAEOTNMLANASTEPARAAGEGFAVVADE 361
Db 212 NVESVAENVEESVAENVEESVAENVEEIVAPTVEIAPTVEIAPTVEIAPTVEIAPTVEI 268
QY 362 VKLAESRQOSTREVELVEQWQAE--TEETVDQUDENQNRIGESVEYV-----EAMET 414
Db 269 VEESVEENVEES--VAENVEESVAENVEESV--AEENVEESVAENVEEIVAPTVEIAPT 324

QY 415 LOEIT--DAVEDAASQOEYSTATDEQAVS---TEEVAENVGDVDRAGEIAALDDIAD 469
Db 325 VEELIAPSVESVAPSVESVEENVEESVAENVEESVAENVEEIVAPTVEIAPTVEI 378
QY 470 ATDOQV--RTVEEY 481
Db 379 PTVEIAPTVEI 392

RESULT 11
US-09-712-363-259
; Sequence 259, Application US/09712363
; Patent No. US20020164588A1
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, David
; APPLICANT: Rotstein, Sergio H.
; APPLICANT: Margotte, Edward M.
; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
; FILE REFERENCE: 07419-032001
; CURRENT APPLICATION NUMBER: US/09/712,363
; PRIOR FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: PCT/US00/02246
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/117,531
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 60/117,844
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 60/118,206
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: 60/126,593
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/134,093
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/134,092
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/165,124
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/165,086
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 292
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 259
; LENGTH: 1289
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-712-363-259

Query Match 7.3%; Score 175.5; DB 9; Length 1289;
Best Local Similarity 24.0%; Pred. No. 0.0027;
Matches 134; Conservative 62; Mismatches 197; Indels 165; Gaps 24;

QY 25 RIG-LDEAEIAMRSLFTGIDDD---TMAALAEOPLFEATADALVTFYDHLSEYENTD 80
Db 524 RIGELDQGEV-----GLDEHHERTVAAI-----RLADEVVAE---LQSAERAA 563
QY 81 LFANSTKTVQAKETQAEYLLGLRGGEYDREYAAORARIKHIDVNLGPDVYLG-----A 136
Db 564 -----ERQVASLRARI-----DALVVGLOQRKDGAAMLA 591
QY 137 YTRYTGLDALADVDVADRGEAA-----AAVDELVARFLPMKLLTFD----- 181
Db 592 HNSGAGLFFSINQIVYVRSGYEALAAALGPAAADALAVDGLTAGSVAALQAOQDGRA 651
QY 182 -----QOIAMDYIDSVQRLHDEIDSRQELANAVATHVEAPLSLEATSOQV 229
Db 652 VLVLSDPAPAPQASAGEMLPQGAQWALDLVESPPQVQAMI---AMLSGV-AAYNDL 706
QY 230 AERTDTMRARTDOQVRMADVSREI-----SSVASVEEYASTADVVRTSEDAAE 280
Db 707 TEAMGLIVEIRP---LRAVTVDGDLVGAGWVSGGSQDRKSLTLEVTSEIDKARSELAAAE 763

QY 281 LAOQGEA-----AADDALATMTDIDEATDGTAGVGEOLGE-----RAADVE 321
 DB 764 LAAQNALMALAGLTQSGARQDAEQALALANESDTA-----ISMRYQDLGLGEAGEAEE 820
 QY 322 -----SVTGYIDI-----AEQTNML-----ALNASTEARAGE 350
 DB 821 WNRLLQORTEGEAVFTQTLTDVLTQLETLRKQAEFTQRYQVAPIDRQALISAADARAGVE 880
 QY 351 AGEFVAVADEKALAEESREGEVVELVEEQMAETEETVQDLQEVNORIGEGVERVE 410
 DB 861 VEARLAVTAEERANAVKGRDLSLRRAAAER---EARYRAQOARARARLHAAVAVAAD 937
 QY 411 AMETLOETITDAVEDAASGMOEVSSTATDEQAVST-EVAEMVDGVDDRAGEIAAL--DDI 467
 DB 938 CGRLLAGRLHRAVDSASQLRASAAQROQRALAAAVRDEVNTLSARVGEILDSLRDEL 997
 QY 468 ADATDOQVTVEEVRETV 485
 DB 998 ANA--QAALRIEQLQOMV 1013

RESULT 12
 US-09-845-583-8
 : Sequence 8, Application US/09845583
 : Patent No. US20020142954A1
 : GENERAL INFORMATION:
 : APPLICANT: Burgeson, Robert
 : APPLICANT: Brunken, William Joseph
 : APPLICANT: Champiland, Marie-France
 : APPLICANT: Hunter, Dale
 : TITLE OF INVENTION: LAMININ 15 AND USES THEREOF
 : FILE REFERENCE: 10287-056001
 : CURRENT APPLICATION NUMBER: US/09/845,583
 : CURRENT FILING DATE: 2001-04-30
 : PRIOR APPLICATION NUMBER: US 60/200,863
 : PRIOR FILING DATE: 2000-05-01
 : NUMBER OF SEQ ID NOS: 18
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 8
 : LENGTH: 1798
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 US-09-845-583-8

Query Match 7.3%; Score 175.5; DB 10; Length 1798;
 Best Local Similarity 20.4%; Pred. No. 0.0039;
 Matches 112; Conservative 80; Mismatches 230; Indels 127; Gaps 17;

QY 7 TLVTADVNGID-----GHALADRIGLDEAETAMRLSFTGID-----DDTMAALAE 53
 DB 1277 TQLEADLTVDQDENENANHALS---GLERDLALNLTLRQDLQDLKHSNPLGAYDSI 1333
 QY 54 QPLFEATADALVTFDYHLESTERTQDLFANSTKTEVQDLKETQAEYLLGLGGEYDTEYA 113
 DB 1334 RHAHSQSAEA-----ERRANTSALAVPSPVNSASARHRTALMDAQEDPNSKIM 1384
 QY 114 AQRATIGI-----HDVL-----GLG----- 129
 DB 1385 ANQRALGKLSAHTHTLSTLDINELVCAOGLHHDRTSPGGGAGCDEDEQPPRCGLSCNG 1444
 QY 130 ----PDVIGATRYTYGTLALAD-----DVAADRGEFAAAYDELVARFLPMLKLTF 180
 DB 1445 AATADLADLGRARHTQAEQRLALAGGSLSNVAETBRQASAQORAAQ-----AL 1495
 QY 181 DOQIAMDTYIDSYAQRLLHDEIDSROELAN-----AVATHEAPLSLSLEATSODVA 230
 DB 1496 DKANASRCOVEQANQELQELIQSVKDFLNOEGADPDSTIEMVATRY-----LELSIPASA 1549
 QY 231 ERTDTMARATDOQVBRMADYSREISSVSASVEAVSTADVVRT---SEDAALAQOQGA 287
 DB 1550 EOIQHLAIAIAERVSRLDVIDVADILARTVGDVRRABQLLODARASMADEQOKAETVOA 1609
 QY 288 AADALATMTDIDEATDGTAGVE-----QLGERADVESVTGYIDIABQTMNMLN 340

DB 1610 ALEEAQRAQIAQOAGAIRGAVADTRDTQTLVQVDERMGAERALSAGERRAQ-----LD 1664
 QY 341 ASTEARAGEAGEFVAVADEKALAEESREGEVVELVEEQMAETEETVQDLQEVNOR 400
 DB 1665 ALLEALKTLKRAGNSLAASAEETAGSAGORAO--FAEQLLRBPILDQYOTVRLAE---R 1719
 QY 401 IGEVY-----ERYEAMETLOETITDAVEDAASGMOEVSSTATDEQAVSTEEVAVMGVDOR 456
 DB 1720 KAQGVLAQAQRAEQLPDEARDLLQAADKLRQLDELECTYEEENERALESKAQOGLQEAR 1779
 QY 457 AGEIAAALD 465
 DB 1780 MRSVLQAIN 1788

RESULT 13
 US-09-815-242-13262
 : Sequence 13262, Application US/09815242
 : Patent No. US20020061569A1
 : GENERAL INFORMATION:
 : APPLICANT: Haselbeck, Robert
 : APPLICANT: Ohlsen, Karl L.
 : APPLICANT: Zyskind, Judith W.
 : APPLICANT: Wall, Daniel
 : APPLICANT: Twilick, John D.
 : APPLICANT: Carr, Grant J.
 : APPLICANT: Yamamoto, Robert T.
 : TITLE OF INVENTION: Identification of Essential Genes In
 : FILE REFERENCE: ELITRA.011A
 : CURRENT APPLICATION NUMBER: US/09/815,242
 : CURRENT FILING DATE: 2001-03-21
 : PRIOR APPLICATION NUMBER: 60/191,078
 : PRIOR FILING DATE: 2000-03-21
 : PRIOR APPLICATION NUMBER: 60/206,848
 : PRIOR FILING DATE: 2000-05-23
 : PRIOR APPLICATION NUMBER: 60/207,727
 : PRIOR FILING DATE: 2000-05-26
 : PRIOR APPLICATION NUMBER: 60/242,578
 : PRIOR FILING DATE: 2000-10-23
 : PRIOR APPLICATION NUMBER: 60/253,625
 : PRIOR FILING DATE: 2000-11-27
 : PRIOR APPLICATION NUMBER: 60/257,931
 : PRIOR FILING DATE: 2000-12-22
 : PRIOR APPLICATION NUMBER: 60/269,308
 : PRIOR FILING DATE: 2001-02-16
 : NUMBER OF SEQ ID NOS: 14110
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 13262
 : LENGTH: 1179
 : TYPE: PRT
 : ORGANISM: Streptococcus pneumoniae
 US-09-815-242-13262

Query Match 7.3%; Score 174; DB 10; Length 1179;
 Best Local Similarity 19.2%; Pred. No. 0.003;
 Matches 113; Conservative 101; Mismatches 201; Indels 174; Gaps 24;

QY 49 AATAQOPLEATADALVTFDYHLESTERTQDLFANSTKTEVQDLKETQAEYLLGLGREGY 108
 DB 457 ATAQDEEQKTSYQAQOSQLEFRLDSLKKQ-----ARAQSLNITLNNHSNFAGV----- 507
 QY 109 DTEVAQARARICKIDHYLG--LGPVYLGATRYTYTGL---LDALADVAVADGEFAAAA 163
 DB 508 -KSVLQEKORIGITIGAYSEHLTFDYY-----YQTALEIALGASSQHIIVEDESATKA 560
 QY 164 VDELY-----ARFLPM--LKLTF-----DDQIAMDTYIDSYAQRLL 197
 DB 561 IDFLKRNRYGRATFLPLTTIKARTISSQONQMAIVSPGFLGMADELVTFDTRLEALFKML 620
 QY 198 -----HDEIDROELANNAVATHEAPLSLSLEATSODVAVERTDTMA-----RTDDQV--- 244

Db 621 LATTATFDYVEFAEARQRYQVR--MTLDTTEL-----RTGGSYGANGARONNSIFIK 674
OY 245 DRMAVDSREISSVASY--DEVA--STADVRRTSDEALAAQGEA-----288
Db 675 PELBOLKEIAADEBASIGSEAAALKTLQDQMAALTEFLERLEIKSOGEGARQIOEGCLATQ 734
OY 289 ----ADALATMTDI--DEATDGVTAQ-----VEQLGERADVESVTGIDIAEQT 334
Db 735 QTSQOYVELETLMKLOEEEDIRLSEGDWQADKEKQCESLATIASSEKQNLAEIEETISNK 794
OY 335 NMLA--LMSIFAARA-----GAGGEPFANVADEVVAL-----365
Db 795 NAIQERYONLQEEVQAQRLKTELQGRYEVADIERLQEDLNLTNEOEIEIOMLOEKY 854
OY 366 -----AESRQSTREVELYEQMOAETEEYDQDDEVQRIQEG-----404
Db 855 DNLEKVTETLLSQOAESEKTKQKTNLQOGLIRKQFELDDIGQDLDDIASHLDQARQNEEM 914
OY 405 -----VERVEAMETLQ-----EITDAVED-----AASQMOEVSTAT 436
Db 915 IRKQTRAEAKKEKYSERLRHLQNLQTDQOISYTEALEKHALENLNLAQEQVQDLEKAI 974
OY 437 DEQAVSTEEVBAWVDGDDPAGELAAALDDIADATDQOQVTVVEERTV 485
Db 975 RSLGPNVLEAIDQEEVHNRLDFLNSQRDILSAKNLLELTITEMNDEV 1023

RESULT 14

US-09-938-275-9
; Sequence 9, Application US/09938275
; Patent No. US20020111309A1
; GENERAL INFORMATION:
; APPLICANT: Gerardo Castillo
; APPLICANT: Alan Snow
; TITLE OF INVENTION: Therapeutic and Diagnostic Applications
; TITLE OF INVENTION: of Lamelin and Lamelin-Derived Protein Fragments
; FILE REFERENCE: PROTEO.P03
; CURRENT APPLICATION NUMBER: US/09/938,275
; CURRENT FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1798
; TYPE: PRT
; ORGANISM: Homo Sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Swissprot P55268
; DATABASE ENTRY DATE: 1996-10-01
US-09-938-275-9

Query Match 7.2%; Score 173.5; DB 10; Length 1798;
Best Local Similarity 20.2%; Pred. No. 0.0051;
Matches 111; Conservative 83; Mismatches 228; Indels 127; Gaps 17;

OY 7 TLVTADVRCID-----GHALADRIQIDEAEIAMRLSFTGID-----DDTMAALAE 53
Db 1277 TQLEADLTQVDQENENANHALS---GLERDLALNTLRLQDHLDLKHSNFGAVDSI 1333
OY 54 QPFEATADALVDFDHLSEYERTQDLFANSKTYEQVLKETQAEYILGIGREYDTEYA 113
Db 1334 RHASQSAEA-----ERRANTSA LAVSPVSNSASARHRTALMDAQKEDFNSKHM 1384
OY 114 AQRARIGK-----IHVUL-----GLG-----129
Db 1385 ANRALKLSAHTHTLSLTDINELVCGAPDCAATSPGCGACRDEGQPPCGGLSCNG 1444
OY 130 ----PVTYIGAVTRYTGLDALAD-----DVADRGEEAAAANVDELVARFLPMLKLTFF 180
Db 1445 AATATADLALGSAHRTQAEIQRALAEGSSILSRVAETRRQASQAQRAQ-----AL 1495
OY 181 DQOIAADTYIDTAQRULHDEIDSRQELAN-----AVATHVEAPLSSLEATSDOVA 230

Db 1496 DKANASRGVEQANQELQELIOSVKDFLNOGADPDSIENAVATRV-----LELSIPASA 1549
OY 231 ERTDTMRATPDQVDRMAVDSREISSVASEEVAESTADVRRRT-----SEDAEALAAQGEA 287
Db 1550 EQIQLHAGALAEVRYSLADVALIARTVGVRRAREQLDQARRARRASAEDEKKALEYVQA 1609
OY 288 AADALATMTDIDEATDGVTAQVE-----QJGERADVESVTGIDIAEQTNMLALN 340
Db 1610 ALBEAQRAQCIAGCAIRGAVADRTRDEQTLTYQJQERMAQERALLSSNGERARQ-----LD 1664
OY 341 ASTEAAARGAGGCFANVADVEKALAEESRQSTREVELYEQMOAETEEYDQDDEYNOR 400
Db 1665 ALLEALKLRAGNSLAATAETRAGSAQGRAQ--EAEQLRGLQDQOYQTVKALAE--R 1719
OY 401 ICGEV-----ERVEAMETLQETDADADAASQMOEVSTATDEQAVSTEEVBAWVDGDDR 456
Db 1720 KAGGVLAQAARARQRLREARDDLQAAQDKLQRLQELQETYEENBRALESNAQDLGEGAR 1779
OY 457 AGEIAMAALD 465
Db 1780 MRSVLQAIN 1788

RESULT 15
US-09-873-676-113
; Sequence 113, Application US/09873676
; Patent No. US20020077289A1
; GENERAL INFORMATION:
; APPLICANT: Macdonald, Nicholas J.
; APPLICANT: Sim, Kim L.
; TITLE OF INVENTION: Angiostatin and Endostatin Binding Proteins and Methods of Use
; FILE REFERENCE: 05213-0378 (43170-259333)
; CURRENT APPLICATION NUMBER: US/09/873,676
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 60/209,065
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: US 60/289,387
; PRIOR FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 113
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-873-676-113

Query Match 7.2%; Score 173; DB 10; Length 1786;

Best Local Similarity 19.7%; Pred. No. 0.0055;
Matches 103; Conservative 92; Mismatches 192; Indels 136; Gaps 20;

OY 45 DTMAALAEQ-----LFEATADALVDFDHLSEYE--RTODLFANSTRKTVQLK--ET 95
Db 1232 DILAQSAAPRLKINIGLFE--EAEKLIKDVTEMAQVEVRLSTTGSNSNTAELEDSLOT 1290
OY 96 QAEYLLGIGREYDTEYAQARARIGKIHVULGAGPDVYLGAVTRYTGLDALADVVAD 155
Db 1291 EAESL-----DMTVKELAEQLEFIKNSDIRGA-----LDSITKYRQMSLEA-----1331
OY 156 RGEAAAVAVDELVARFLPMKLTLPDQOIAMQVYIYSAQRULHDEIDSRQELANAVATV 215
Db 1332 -EEVNASSTTEPNS-----TVQSALMRDRVEDVMMERBSQKEKQEBQARLDEL 1381
OY 216 EAPLSLEATSDVART-----DTMRARTDD-----242
Db 1382 AGKLOSLDLSA--AAEMTCTPPGASCSSECEGPNKRTIDEGGRKCGPGCGGLVYVAHN 1439
OY 243 QVDRMAVDSREISSVASEEVAESTADVARTSEDAEALAAQGEAADAAL-----ATMTD 298
Db 1440 AMQKAMDLDQDVLASALAEVQLS-----KNVSEAKIRADEAKQSAEDILLTKTNATKER 1492
OY 299 IDEATDGVTAQVQ-----LGERADVESVTGV-----IDDIAEQTN 335
Db 1493 MDSNEELRNLIKQIRNFLTQDSADLDSIEAVANVELKMEMPSTPOOLQNLMTEDIRERVE 1552

Oy 336 MLALNASI-----EARAGEGEGFAVVADVEVKALAEESREOSTRVEEL 379
| : | : | : | : | : | :
Db 1553 SLSQVEVILQHSNADIRAREMLLEPAKRKSKSTDVKTADMKKEALEEHEKQVAAEKA 1612
Oy 380 VEOMQAEETEETVDQLEDEVNORIGEGVERVEAMEETLOEITDAVEDAASGMOEVSTATD-- 437
: : : : : : : : : : : : : : : :
Db 1613 IKQADEDIGTONLTSISEETAASEETLFNNSQRISELERNVEELKRAKAQNSGEAEYI 1672
Oy 438 EQAV-----STEEVAEMVDC-VDDRAGE----IAAALDDIADA 470
| : | : | : | : | : | :
Db 1673 EKVVYTVKQSAEDVKKTLDGELDEKYYKVENLIAKKTESADA 1715

Search completed: January 2, 2003, 12:40:02
Job time : 21 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2003, 12:36:16 ; Search time 96 Seconds
(without alignments) 678.745 Million cell updates/sec

Title: US-09-455-978B-2

Perfect score: 2394

Sequence: 1 MSMDNDLVADVRNGIDGH.....ATDQVRFVEEVRNWKLS 489

Scoring table:

BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: A_Geneseq_101002.*
2: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
3: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
4: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
5: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
6: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
7: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
8: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
9: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
10: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
11: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
12: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
13: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
14: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
15: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
16: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
17: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
18: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
19: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
20: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
21: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
22: /SID22/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
23: /SID22/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
24: /SID22/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2394	100.0	489	22	AAE04641
2	412.5	17.2	432	22	AAE04642
3	387.5	16.2	739	22	AAE04643
4	380	15.9	435	22	AAE04643
5	375	15.7	501	22	AAE04643
6	359.5	15.0	891	21	AAE04643
7	356.5	14.9	374	22	AAE04643
8	339.5	14.2	601	23	AAE04643
9	298	12.4	511	21	AAE04643
10	297.5	12.4	1137	22	AAE04643

11	289	12.1	565	19	AAE04643
12	289	12.1	565	19	AAE04643
13	289	12.1	565	19	AAE04643
14	284	11.9	630	22	AAE04643
15	284	11.9	630	22	AAE04643
16	284	11.9	630	22	AAE04643
17	280	11.7	675	19	AAE04643
18	276	11.5	293	18	AAE04643
19	273.5	11.4	664	18	AAE04643
20	269	11.2	723	20	AAE04643
21	269	11.2	753	20	AAE04643
22	266	11.1	606	20	AAE04643
23	266	11.1	633	20	AAE04643
24	257	10.7	431	21	AAE04643
25	254	10.6	433	17	AAE04643
26	253	10.6	300	23	AAE04643
27	252.5	10.5	438	18	AAE04643
28	249.5	10.4	701	22	AAE04643
29	242.5	10.1	191	22	AAE04643
30	221	9.2	566	22	AAE04643
31	214.5	9.0	110	23	AAE04643
32	207.5	8.7	524	20	AAE04643
33	207.5	8.7	553	20	AAE04643
34	203	8.5	39	22	AAE04643
35	202.5	8.5	955	15	AAE04643
36	202.5	8.5	955	17	AAE04643
37	199	8.3	258	22	AAE04643
38	198	8.3	98	21	AAE04643
39	196.5	8.2	1857	23	AAE04643
40	194.5	8.1	1851	22	AAE04643
41	194.5	8.1	1960	22	AAE04643
42	194.5	8.1	2143	22	AAE04643
43	194	8.1	1786	18	AAE04643
44	194	8.1	1787	23	AAE04643
45	194	8.1	3854	22	AAE04643

ALIGNMENTS

RESULT 1	AAE04641	AAE04641 standard; Protein: 489 AA.
AC	AAE04641:	
XX		
XX		
DT	04-SEP-2001 (first entry)	
XX		
DE	Halobacterium salinarum HemAT-Hs protein.	
XX		
KW	Haem binding protein; HemAT-Hs; HemAT-Bs; gaseous ligand sensor; oxygen storage; artificial photosynthesis; signalling function;	
KW	alpha-haemoglobin; myoglobin; therapy.	
XX		
OS	Halobacterium salinarum.	
XX		
PN	W0200140475-A2.	
XX		
PD	07-JUN-2001.	
XX		
PF	05-DEC-2000; 2000WO-US33048.	
XX		
PR	06-DEC-1999; 99US-0455978.	
XX		
PA	(UYHA-) UNIV HAWAII.	
XX		
PI	Alam M, Larsen R;	
XX		
DR	WPI: 2001-374832/39.	
XX		
DR	N-PSDB; AAD08991.	
XX		
PT	Novel isolated bacterial heme binding protein, HemAT-Hs or HemAT-Bs which reversibly binds oxygen with low affinity, useful for controlled	

PT storage of oxygen and for sensing gaseous ligands such as oxygen
XX Claim 6; Page 10; 94pp; English.
XX
CC The present invention relates to isolated archaeal and bacterial haem
CC binding protein, Hemat-Hs or Hemat-Bs which reversibly binds oxygen
CC with low affinity. Haem binding protein is useful for controlled storage
CC of oxygen by allowing haem binding protein to bind and store oxygen, and
CC triggering the release of oxygen from haem binding protein by activating
CC the signalling domain. Haem binding protein is useful for sensing gaseous
CC ligands such as oxygen, NO, CO, or CN. Blood substitute comprising haem
CC binding protein is useful for creating a patient suffering from low
CC blood levels by administering and regulating the oxygen binding of the
CC haem-binding protein by modifying the signalling domain. Haem binding
CC protein is useful for haem-based catalysts, for artificial
CC photosynthesis and for identifying potential signalling functions of
CC mutated alpha-haemoglobin and myoglobin causing several diseases.
CC The present sequence is Halobacterium salinarum Hemat-Hs protein which
CC is salt tolerant.
XX
SQ Sequence 489 AA;
Query Match 100.0%; Score 2394; DB 22; Length 489;
Best Local Similarity 100.0%; Pred. No. 1.8e-168;
Matches 489; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MSNDNDLTADYRNGIDGHALDRIGDEAEIAMRLSFTGIDDTMAALAEQPLEAT 60
DB 1 MSNDNDLTADYRNGIDGHALDRIGDEAEIAMRLSFTGIDDTMAALAEQPLEAT 60
OY 61 ADALVDFDHESEYRTODLFANSTKTYEQLEKTAQEVLLGSGREYDTEVAQAARIG 120
DB 61 ADALVDFDHESEYRTODLFANSTKTYEQLEKTAQEVLLGSGREYDTEVAQAARIG 120
OY 121 KIHNVGLGPDVYLGATRYTGGLLDALADDDVADRGEEAAAIVDELVARFLPKLTF 180
DB 121 KIHNVGLGPDVYLGATRYTGGLLDALADDDVADRGEEAAAIVDELVARFLPKLTF 180
OY 181 DOOIAMDTYDSYQRLHDEIDSRQELANNAVATNVEAPLSLEATSQDAERTDTMARAT 240
DB 181 DOOIAMDTYDSYQRLHDEIDSRQELANNAVATNVEAPLSLEATSQDAERTDTMARAT 240
OY 241 DDOVDMADVSRFTSSVSASVEEVASTDVVRTSEDAELAQQGAAADALATMTDID 300
DB 241 DDOVDMADVSRFTSSVSASVEEVASTDVVRTSEDAELAQQGAAADALATMTDID 300
OY 301 EATDGTAGVEQGERAAVDESVTGVIDIAEOTNMLALNASTEARAGGEGFAVVAD 360
DB 301 EATDGTAGVEQGERAAVDESVTGVIDIAEOTNMLALNASTEARAGGEGFAVVAD 360
OY 361 EYKALAEESREOSTRVEELVEOMQAEETVDQLDDEVNORIGRGVERVEAMETLOEITD 420
DB 361 EYKALAEESREOSTRVEELVEOMQAEETVDQLDDEVNORIGRGVERVEAMETLOEITD 420
OY 421 AVDDAASGQEVSTATDEQAVSTEEVAENVDGVDPRAGETAAALDIADATDOQVTVEE 480
DB 421 AVDDAASGQEVSTATDEQAVSTEEVAENVDGVDPRAGETAAALDIADATDOQVTVEE 480
OY 481 VRETGKLS 489
DB 481 VRETGKLS 489

KW Haem binding protein; Hemat-Hs; Hemat-Bs; gaseous ligand sensor;
KW oxygen storage; artificial photosynthesis; signalling function;
KW alpha-haemoglobin; myoglobin; therapy.
XX
XX Bacillus subtilis.
OS
PN WO200140475-A2.
XX
XX 07-JUN-2001.
XX
XX 05-DEC-2000; 2000WO-US33048.
XX
XX 06-DEC-1999; 99US-0455978.
XX
XX (UYHA-) UNIV HAWAII.
XX
XX Alam M, Larsen R;
XX
XX WPI; 2001-374832/39.
XX
XX N-PSDB; AAD08992.
DR
PT Novel isolated bacterial heme binding protein, Hemat-Hs or Hemat-Bs
PT which reversibly binds oxygen with low affinity, useful for controlled
PT storage of oxygen and for sensing gaseous ligands such as oxygen
XX
XX Claim 8; Page 11; 94pp; English.
XX
CC The present invention relates to isolated archaeal and bacterial haem
CC binding protein, Hemat-Hs or Hemat-Bs which reversibly binds oxygen
CC with low affinity. Haem binding protein is useful for controlled storage
CC of oxygen by allowing haem binding protein to bind and store oxygen, and
CC triggering the release of oxygen from haem binding protein by activating
CC the signalling domain. Haem binding protein is useful for sensing gaseous
CC ligands such as oxygen, NO, CO, or CN. Blood substitute comprising haem
CC binding protein is useful for treating a patient suffering from low
CC blood levels by administering and regulating the oxygen binding of the
CC haem-binding protein by modifying the signalling domain. Haem binding
CC protein is useful for haem-based catalysts, for artificial
CC photosynthesis and for identifying potential signalling functions of
CC mutated alpha-haemoglobin and myoglobin causing several diseases.
CC The present sequence is Bacillus subtilis Hemat-Bs protein.
XX
SQ Sequence 432 AA;
Query Match 17.2%; Score 412.5; DB 22; Length 432;
Best Local Similarity 26.2%; Pred. No. 2.4e-22;
Matches 114; Conservative 85; Mismatches 187; Indels 49; Gaps 8;
OY 31 AETAMRLSTGIDDTMAALAEQPLEATADALVDFDHESEYRTODLFANSTKTYE 90
DB 31 AETAMRLSTGIDDTMAALAEQPLEATADALVDFDHESEYRTODLFANSTKTYE 90
OY 91 OKETQAEVLGSGREYDTEVAQAARIGKIHNVGLGPDVYLGATRYTGGLLDALAD 150
DB 91 OKETQAEVLGSGREYDTEVAQAARIGKIHNVGLGPDVYLGATRYTGGLLDALAD 150
OY 151 DVVADRGEEAAAIVDELVARFLPKLTFDQOIAADTYDSYQRLHDEIDSRQELANA 210
DB 149 -----EASITNQOELLAKATKILNLEQGLVLEAPOSEVNG--TRDQGEKKLHLH- 199
OY 211 VATNVEAPLSLEATSQDAERTDTMARATDDOVDPMADVSRFTSSVSASVEEVASTDAD 270
DB 200 -----OKIOETSGISA---NLFSETSRSVOELVDKSEGISQASKAGTVTSSTVEE 246
OY 271 VRTSEDAELAQQGAAADALATMTDIDENATDGTAGVEQGERAAVDESVTGVIDI 330
DB 247 -KSIGGRKELEVQOQO-----NNKIDTSLVQIEKEVAKIDETAQOIEKIFGLVTGI 296
OY 331 AEOITNMLALNASTEARAGGEGFAVVADDEVKALAEESREOSTRVEELVEOMQAEET 390
DB 297 AEOITNMLALNASTEARAGGEGFAVVADDEVKALAEESREOSTRVEELVEOMQAEET 390
OY 391 VDQLDDEVNORIGRGVERVEAMETLOEITDAVEDA-----ASGQEVSTAT 436

Db 357 SKHIDVNEVSESEKMTQINRLFDEIVHSKISKESQKIDVDLQAFGLGDEVSRAV 416
Oy 437 DEQAVSTEEVAEMVD 451
Db 417 SHVAASVDLVLTE 431

RESULT 3
AAB96493
ID AAB96493 standard; Protein: 739 AA.
XX
AC AAB96493:
XX
DT 29-OCT-2001 (first entry)
XX
DE Putative sensory transduction histidine kinase and response regulator #3.
XX
KW Hyperthermophilic archaeon; hyperthermophilic protein.
XX
OS Pyrococcus abyssi.
XX
PN FR2792651-A1.
XX
PD 27-OCT-2000.
XX
PE 21-APR-1999; 99FR-0005034.
XX
PR 21-APR-1999; 99FR-0005034.
XX
PA (CNRS) CNRS CENT NAT RECH SCT.
XX (IFRE-) IFREMER INST FR RECH EXPL MER.
XX
PI Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;
PI Querellou J, Weissenbach J, Saurin W, Hellig R;
XX WPI: 2001-126236/14.
XX
DR New nucleotide sequences isolated from Pyrococcus abyssi encode
XX proteins useful in industry -
PT
PS Claim 7; Pages 1203-1205; 1657pp; French.
XX
XX The present invention relates to the genomic sequence of Pyrococcus
CC abyssi (see AAF66431 and AAH41223-7) and P. abyssi proteins. P. abyssi is
CC a hyperthermophilic archaeon, which is isolated from deep-sea
CC hydrothermal vents. The present sequence is one such P. abyssi protein.
CC The proteins of the present invention have various potential industrial
CC uses, since the proteins are stable at very high temperatures, some up to
CC 110 degrees centigrade.
CC Note: This patent is in the same patent family as WO200065062, which
CC contains additional sequences as shown in AAB99132-AAB99143,
CC AAH75903-AAH75920 and AAG66436.
CC
XX
SQ Sequence 739 AA:

Query Match 16.2%; Score 387.5; DB 22; Length 739;
Best Local Similarity 28.6%; Pred. No. 3.4e-20;
Matches 126; Conservative 83; Mismatches 150; Indels 81; Gaps 15;

Oy 83 ANSTFVEQLKETQAEYLIGRGEDYEYAAQ-----RARGKI-----HDVILGIG 129
Db 331 SNTLAPLEKLR--YAAQALAEGRKQVSEYELKQIRYLERDELGALIQAFEAVSKLVGT- 387
Oy 130 PDVYLGAVTRYTGLDLADLAD-----VADRGEEAAAVDELVARFLPMILKLTLPDQ 183
Db 388 ----LNAISK-----KLERLAEGDLSNGLTVEYRG-----LROIIE-- 421

Oy 184 IAMDYIDSYAQRHLDEIDSRQELANAVATHEAPLSSLEATSDVAERTDMARKTDQ 243
Db 422 -----ISVYETFPRESIGSLVEMAN---DLEKRANALAOVSKDVTAINQVN-EAIIQ 470
Oy 244 VDRMADVSR-ISSVASVEEVASTADVRRTE-----DAELAAQGEAAADALAT 295

Db 471 VSIENQROQERINNETDGMRLVAQSESVRAMEFSCAVTEEVYSIANEGSQGDEALKR 530
Oy 296 MTDIDEATDGYTAGVEOLGERADVESYGIIDYIAEOTNNALALNASTEAPRAGAGCGF 355
Db 531 IEDIOHMMRSRIETEVSVKAEVMSRNIEETVNT\$IAEOTNNALALNASTEAPRAGAGCGF 590
Oy 356 AVVADEVKALAESESESTRAVELVEQMAETETVDQDDEVNORIGGEVREEMETL 415
Db 591 AVVAQEIIRKLAESKQADNITSIDKITDELKAEVATK-----EGSVYIGSETL 643
Oy 416 QEITDVEDAASGQEV\$---TATDEQAVSTEEVAE---MYDGVDRAGEIAALDDIA 468
Db 644 RDTIGYLANIATILLQETSERMTTVYEQIVRQGEYDKALRALENLAASAEETASAEVS 703
Oy 469 DATDQQVTVBEVRETVGKL 488
Db 704 SAIEQOTAAIEELRRRAOEL 723

RESULT 4
AAB96483
ID AAB96483 standard; Protein: 435 AA.
XX
AC AAB96483:
XX
DT 29-OCT-2001 (first entry)
XX
DE Putative sensory transduction histidine kinase and response regulator #2.
XX
KW Hyperthermophilic archaeon; hyperthermophilic protein.
XX
OS Pyrococcus abyssi.
XX
PN FR2792651-A1.
XX
PD 27-OCT-2000.
XX
PE 21-APR-1999; 99FR-0005034.
XX
PR 21-APR-1999; 99FR-0005034.
XX
PA (CNRS) CNRS CENT NAT RECH SCT.
XX (IFRE-) IFREMER INST FR RECH EXPL MER.
XX
PI Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;
PI Querellou J, Weissenbach J, Saurin W, Hellig R;
XX WPI: 2001-126236/14.
XX
DR New nucleotide sequences isolated from Pyrococcus abyssi encode
XX proteins useful in industry -
PT
PS Claim 7; Pages 1189-1190; 1657pp; French.
XX
XX The present invention relates to the genomic sequence of Pyrococcus
CC abyssi (see AAF66431 and AAH41223-7) and P. abyssi proteins. P. abyssi is
CC a hyperthermophilic archaeon, which is isolated from deep-sea
CC hydrothermal vents. The present sequence is one such P. abyssi protein.
CC The proteins of the present invention have various potential industrial
CC uses, since the proteins are stable at very high temperatures, some up to
CC 110 degrees centigrade.
CC Note: This patent is in the same patent family as WO200065062, which
CC contains additional sequences as shown in AAB99132-AAB99143,
CC AAH75903-AAH75920 and AAG66436.
CC
XX
SQ Sequence 435 AA:

Query Match 15.9%; Score 380; DB 22; Length 435;
Best Local Similarity 27.5%; Pred. No. 6.1e-20;
Matches 109; Conservative 91; Mismatches 129; Indels 68; Gaps 12;

Oy 107 EYDFEYAAQRIRIGIHVILGIPDVYLGATRYTGLDLADVDVADRGEEAAAVDE 166

Db 63 EYKQLOREIDEIVKILIDRIAQ6-DL-----SYEDKREIGMGCAVRGIGED 107
QY 167 LVARFLPMLKLTFEDQOIAMDFYIDSYAORLHDEIDSRQELAAV-ATHVEAPLSLEAT 225
Db 108 L-RKSIQVLYNKKNAIDVRNHTTKVMEKIEQVADSVQOVAFINQVSTEA-----QRE 161
QY 226 SODVARTDTMRATDDQVDRMADVSRISVSASVEEVASTADVTRTSEDAEALAQOQ 285
Db 162 QENISMTDTMR-----YIHDSKE--TVSTMEFEAS-----MREMAQLAKEGEGK 207
QY 286 EAAADALATMDIDATGCTAGVQOLGERADVESVGVDDIAEQTNMLALNASTEA 345
Db 208 RQADQD-IEIISRMEMKIEETVGVAVEMGK--SININIVISSISEQTNLLALNAIEA 263
QY 346 ARAGEGEGFAYVADENKLAESRQSTRVEELVEOMQAEETBEVDQDDEVNORIGEGY 405
Db 264 ARAGEGKGFAYVADIRKLAESKKAADIRLINO-----IGDKIGESY 309
QY 406 ERVEAEMTLQETDAVEDAASGMEVSTATDEQAVSTEEVAMV-----D 451
Db 310 EYVQGAEEVVKSTEVKESVSYLTQVAKMEMEYKASLREKVIQEGKIEGLRFL 369
QY 452 GVDNRAGETAAADIDATDQOVRTVEEYREYVKL 488
Db 370 NLASAEEYTAAEVSAAEQSSALQELRESVKEL 406
RESULT 5
AAB96709
ID AAB96709 standard; Protein; 501 AA.
XX AAB96709;
AC AAB96709;
XX
DT 29-OCT-2001 (first entry)
XX
DE Putative sensory transduction histidine kinase & response regulator #4.
XX
KW Hyperthermophilic archaeon; hyperthermophilic protein.
XX
OS Pyrococcus abyssi.
XX
PN FR2792651-A1.
XX
PD 27-OCT-2000.
XX
PF 21-APR-1999; 99FR-0005034.
XX
PR 21-APR-1999; 99FR-0005034.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
XX (IFRE-) IFREMER INST FR RECH EXPL MER.
XX
PI Forterre P, Thiery JC, Prieur D, Dietrich J, Lecompte O;
XX Querellou J, Weissenbach J, Saurin W, Heilig R;
XX WPI: 2001-126236/14.
DR
XX
XX New nucleotide sequences isolated from Pyrococcus abyssi encode
PT proteins useful in industry -
XX
XX Claim 7; Pages 1469-1471; 1657pp; French.
PS
XX
XX The present invention relates to the genomic sequence of Pyrococcus
CC abyssi (see AAF86431 and AAH41223-7) and P. abyssi proteins. P. abyssi is
CC a hyperthermophilic archaeon, which is isolated from deep-sea
CC hydrothermal vents. The present sequence is one such P. abyssi protein.
CC The proteins of the present invention have various potential industrial
CC uses, since the proteins are stable at very high temperatures, some up to
CC 110 degrees centigrade.
CC Note: This patent is in the same patent family as WO2000065062, which
CC contains additional sequences as shown in AAB99132-AAB99143,
CC AAH75903-AAH75920 and AAG66436.

XX SQ Sequence 501 AA;
Query Match 15.7%; Score 375; DB 22; Length 501;
Best Local Similarity 24.4%; Pred No. 1.7e-19;
Matches 116; Conservative 111; Mismatches 175; Indels 74; Gaps 12;
QY 65 VDFYDHL-----SYERTODLFANSRTYVQKLETOEAYLLGLGRGEYDEVAA 114
Db 30 ISLDVNLKTKKVVQGTGSVEAIRRELENAK---SQLNET-----LWLSIGVMSI----- 77
QY 115 QRARIKTHDVLGLGPDVYLGAITRYTGLDALADVDVADGEEAANAVIDELVARFLPM 174
Db 78 ----VAVIGALGM-----RLNSTMRPINEMKIAESIEGKISRAREVYSKIYREDE 129
QY 175 L-KLTFDQOIAMDFY--IDSYAORLHDEIDSRQELAAVATHV-----EAPLSLEATSQ 227
Db 130 IGLKIGFPAISODVLYQTEIETERN--EKISEGVSDCLKYHAKGDFESILNSMKITIS 187
QY 228 DVAERTDTMR-----ARTDDQVDRMADVSRISVSASVEEA----- 265
Db 188 NIRELKTIVRDLALTLFSRANDLTRISSEISSEAINQVABAIOQVSVEAQRQENITEIME 247
QY 266 ----STADVTRTSEDAE-----ALAQGEAADALATMTDIDATGCTAGVQO 312
Db 248 GNNITADVTQRTVDAMEEFSGVVNEVLSIAREGKQGEKAIQVEDIDQAMKVITQAVGE 307
QY 313 LGERADVESVGVDDIAEQTNMLALNASTEAARAGEGEGFAYVADENKLAESRQ 372
Db 308 VAEKSNVGDITINADINADQTNLLALNAIEARAGELGRGAVVAOEVNRNLAESKKA 367
QY 373 STRVEELVQOMQAEETBEVDQDDEVNORIGEGYERVEAEMTLQETDAVEDAASGMEY 432
Db 368 AEKIRGILNEIOEKYKAVEETEKGVVDDSVDFLKETVGYLMNIGELLDVESKLDI 427
QY 433 STATDEQAVSTEEVAMVQGVDDRAGEIAADIDATDQOVRTVEEYREYVKL 488
Db 428 KNELANTOEHVENAKKALENLASAQETTASAEVSAQEOASSMEYKRNITEL 483
RESULT 6
AAB26596
ID AAB26596 standard; Protein; 891 AA.
XX AAB26596;
AC AAB26596;
XX
DT 01-FEB-2001 (first entry)
XX
DE Synecocystis sp phytochrome-related gene Cph5.
XX
KW Phytochrome; fluorescent apophytochrome; bilin; Cph.
XX
OS Synecocystis sp.
XX
PN WO200056355-A1.
XX
PD 28-SEP-2000.
XX
PF 14-MAR-2000; 2000MO-US06607.
XX
PR 19-MAR-1999; 99US-0272809.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Lagarias JC;
XX
XX WPI: 2000-602195/57.
DR
XX
XX Fluorescent apophytochrome-bilin conjugates, e.g. Cph2 (derived from
PT Synecocystis species) and phycoerythrobilin conjugate, useful as
XX fluorescent markers for biological research -
XX
XX Example 1; Page 48; 52pp; English.
PS

XX The present invention is related to fluorescent apophytochrome-bilin
CC conjugates, known as phytofluors. An apoprotein known as Cpn2 from
CC *Synechocystis* species is used as the apophytochrome and the bilin is
CC preferably phycoerythrobilin. The phytofluors are useful as fluorescent
CC markers for biological research. The phytofluors have a long wavelength
CC absorption maximum, a high molar absorption coefficient and the
CC recombinant apoproteins can spontaneously assemble with a variety of
CC bilin chromophore precursors. The present sequence is a phytochrome
CC related protein from *Synechocystis* sp.
XX
SQ Sequence 891 AA;
Query Match 15.0%; Score 359.5; DB 21; Length 891;
Best Local Similarity 27.1%; Pred. No. 5,1e-18;
Matches 136; Conservative 86; Mismatches 197; Indels 83; Gaps 19;
OY 23 ADRIGIDEAEIMRLSFTGIDDTMAALAEQP--LFEATADALVTDFDHESEY----- 75
DB 418 ADRV-----IYRRDATAGTIVESVAGYPKALGATADPCFADSY--VEKYSGR 469
OY 76 ERTODLFANSTKT---VEOLK--ETQAEYLGLGGEYDTEYAAORARIGKHIDVLG-- 128
DB 470 QATRDYI-NAGLTPCHIGQLKPEFVKANLY-----APINYGKMLIGLII 512
OY 129 -----GP-----DYVLGATRYTYTGLL-----DALADVVADRGEEAANAADVIAARELP 173
DB 513 AHOCGPRDWHONEIDLEGQLTVQVGLAERSDLAOKIAE--VEQROMREKMRKALE 570
OY 174 MLKLLTFDQOIAMDTYIDSYAORLHDEIDSRQELANAVATHEAPLSLEATSDVAERT 233
DB 571 L--LMEVDPVSRGDLTIRAHV--TDEIGTIDSYNATIESLRITVQVQTASQETET 626
OY 234 DT-----MRRTDQDVRADVSREISSVSASVEEVASTDVDRRSEDAEALAQCE 286
DB 627 DTNEVAVRLOAQANRQALDVAELERLOAMKNSIQAVENMAQAESAVARATQTVDOCE 686
OY 287 AAADALATMTDIDEATGCTAGVBOLEGRADVESYTGVIDDIAEOTMMLANSTIEA 346
DB 687 DAMNTVDGIVAIRETVAATAKQVRKLGESSOKISKVYNLIGSFADQTNLLALNAITEA 746
OY 347 RAGEGGEFAVVADEVKALAEESRQSTREVELVEQMAETEETVDOLDEVNQRIGEGYE 406
DB 747 HAGEGGRGFAVVADEVRSLARQSAEATATIAQLVATIQETNEVYANMAGTEQVYVGRK 806
OY 407 RVEEMETIQEITDAVEAASGMQF-VSTATDEQAVSTEE-----VAEMVD----- 451
DB 807 LVEETRRSLNQIT-AVSAQISGLVEATISMAIESQTSSESVTQTMALVAQIADKNSSEAS 865
OY 452 GVDNRAGEITAAALDDIADATDQ 473
DB 866 GVSATFKELLAVASQSLQEAIVKQ 887
RESULT 7
AAB96382
ID AAB96382 standard; Protein; 374 AA.
XX AAB96382;
AC AAB96382;
XX
DT 29-OCT-2001 (first entry)
XX
DE Putative P. abyssal sensory transduction histidine kinase.
XX
KM Hyperthermophilic archaeon; hyperthermophilic protein.
XX
OS *Pyrococcus abyssi*.
XX
PN FR2792651-A1.
XX
PD 27-OCT-2000.
XX
PF 21-APR-1999; 99FR-0005034.

XX
PR 21-APR-1999; 99FR-0005034.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
PA (IFRE-) IFREMER INST FR RECH EXPL MER.
XX
PI Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;
PI Querellou J, Weissenbach J, Saurin W, Hellig R;
XX
DR WPI; 2001-126236/14.
XX
PT New nucleotide sequences isolated from *Pyrococcus abyssi* encode
PT proteins useful in industry -
XX
PS Claim 7; Pages 1066-1067; 1657pp; French.
XX
CC The present invention relates to the genomic sequence of *Pyrococcus*
CC *abyssi* (see AAF86431 and AAH41223-7) and P. *abyssi* proteins. P. *abyssi* is
CC a hyperthermophilic archaeon, which is isolated from deep-sea
CC hydrothermal vents. The present sequence is one such P. *abyssi* protein.
CC The proteins of the present invention have various potential industrial
CC uses, since the proteins are stable at very high temperatures, some up to
CC 110 degrees centigrade.
CC Note: This patent is in the same patent family as WO200065062, which
CC contains additional sequences as shown in AAB99132-AAB99143,
CC AAH75903-AAH75920 and AAG66436.
XX
SQ Sequence 374 AA;
Query Match 14.9%; Score 356.5; DB 22; Length 374;
Best Local Similarity 26.6%; Pred. No. 2,7e-18;
Matches 105; Conservative 85; Mismatches 145; Indels 59; Gaps 10;
OY 104 GGEED-TEYAAORARIGKHVL-GLGPDVYIGATRYTYGILDALADVVADEGEA 161
DB 19 GKDYENSSNVTSSQAIINLRLDVGKEPQIPG-----LSKEDEE-- 59
OY 162 AAVDELVAFLMLKLTFTDQOIAMDTYIDSYAORLHDEIDSRQELANAVATHEAPLS 221
DB 60 -----VLKVAERLRKGGCKTIN--VKDKIENLKE---VIENLEKIG- 97
OY 222 LEATSDVAERTDTMRARDDQVDRMADVRSISSVSASVEEVASTDVDRRTSEDAEAL 281
DB 98 -EVKGLDGE-VNELVSRNNENVKIAEYNDIQITSAQIEEMNNOAQOLSPALESAM 155
OY 282 AQQGEAAADALATMTDIDEATDGYTAGVEQLERRADVSTGVIDDIAEOTMMLALNA 341
DB 156 AEKGRQISDNVALKYSRISETSREMSDAVRIILEVSKKINDIYVYISIASQTNLLALNA 215
OY 342 STEAARAGGAGGFVNADEVKALAEESRQSTREVELVEQMAETEETVDOLDEVNQR 401
DB 216 STEAARAGGAGGFVNADEVKALAEESRQSTREVELVEQMAETEETVDOLDEVNQR 401
OY 402 GEGVERVEEMETIQEITDAVED-----AASGMQEVSTATDEQAVSTEEVAEMVD 454
DB 269 QENVRTTEVKAQIQLIAFPDIARANETANMKELSEGDIEQANSQMLVDRDISIS 328
OY 455 DRAGETAAALDDIADATDQAVRVEEVRVETVGL 488
DB 329 KDVSQMLNFATQLTDTISGLEKLENEVKEIRTL 362
RESULT 8
AAB49909
ID AAB49909 standard; Protein; 601 AA.
XX AAB49909;
AC AAB49909;
XX
DT 05-FEB-2002 (first entry)
XX
DE -*Listeria monocytogenes* protein #2613.
XX
KM Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;

[illegible]

```

Db      280  SGNVSISSILISAMTTEMIDIEKMMGEQIQIQIGSKKESTDAAMDQGTGGVQNVAAEALV 339
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      272  RRTSEDAEALAOQGEAADALATMTDIDEATDGTAGVEQIGERADVESTGVIDDIA 331
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      340  SDLTRDSAKETNDGIIVINLVSOHDIRISGVKRSSTDVVSQLVNRVGEVEKALDPTVNTIA 399
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      332  EITNNLALNAASIEARAGAGEGFAVADEVALAESREQSTRVLELVEQMAETEERY 391
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      400  DDTNLALNAALIESARAGHGRCGFVAVADEVRKLABOSHLAVVDITVTKKIQTESKTI 459
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      392  DQLEDEVNORIGGEVEREBEMETLQBITDAVEDPASGMOEVSATDQAVSTEEVAMVD 451
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      460  EVMNTGLSESEAGQKITISTEATFTDLNRVNDISQMOVSEFTBEMAGIEEVTSTIS 519
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      452  GY---DDRAGETAAALDDIADATDQOQRYVEEVRVYGLS 489
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      520  DVEISNIGEKSTALFEAVNMKMKVDLVYISEMKRIS 560
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
AAB14148
ID      AAB14148 standard; Protein; 511 AA.
XX
XX      AAB14148;
XX
XX      02-FEB-2001 (first entry)
XX
XX      Bordetella pertussis protein # 2.
XX
XX      Type III secretion system; virulence factor; pathogenicity island.
XX
XX      Bordetella pertussis.
XX
XX      WO200037493-A2.
XX
XX      29-JUN-2000.
XX
XX      21-DEC-1999; 99WO-EP10297.
XX
XX      21-DEC-1998; 98GB-0028217.
XX
XX      (ULBR ) UNIV LIBRE BRUXELLES.
XX
XX      Bollen A, Fauconnier A, Godfroid E;
XX
XX      WPI: 2000-452178/39.
XX
XX      N-PSDB; AAA64886, AAA64890.
XX
XX      Novel polypeptides derived from Bordetella pertussis, useful for
XX      treating and diagnosing Bordetella Infection -
XX
XX      Disclosure: Pages 156-157; 165pp; English.
XX
XX      Bordetella pertussis possesses a type III secretion system. Type III
XX      secretion systems allow bacteria to target virulence factors directly at
XX      host cells. The present sequence is a protein of B. pertussis. No name
XX      has been specified for the present sequence. The coding sequence of the
XX      present sequence is homologous to housekeeping genes of other species,
XX      and so the present protein may have a role in cellular housekeeping. A
XX      pathogenicity island is a compact, distinct genetic unit carrying
XX      virulence genes. The coding sequence of the present sequence is located
XX      within a pathogenicity island (see AAA64890) which also carries, a
XX      number of genes encoding proteins involved in the type III secretion
XX      system of B. pertussis. See AAA64849-A64884 and AAB14111-B14146 for
XX      details of the coding sequences and proteins identified in the
XX      pathogenicity island, of the present invention.
XX
XX      Sequence 511 AA;
XX

Query Match 12.4%; Score 298; DB 21; Length 511;
Best Local Similarity 24.0%; Pred. No. 8.4e-14;
Matches 118; Conservative 87; Mismatches 184; Indels 102; Gaps 15

```

```
OY 33 IAMR-----LSTGIDDDTMAALAEQPLPEATADA---LVTDEYDHL-----73
XX 33 IAMR-----LSTGIDDDTMAALAEQPLPEATADA---LVTDEYDHL-----73
DB 28 LTRWAGSAAEIEGLNOVAANQNDPLFEASGAQORQATOFORYVDPKPEPAAELAAT 87
OY 74 -----SYERTODLFANSTKVEOLKETQAEYLLGLGEGVDETEVAARIGKTHDYLG 127
XX 74 -----SYERTODLFANSTKVEOLKETQAEYLLGLGEGVDETEVAARIGKTHDYLG 127
DB 88 LOTRBAVQSVLDELAA-----VDAGQAEPLAAMHRAQQAHEHAFQD-----131
OY 128 LGPDVYLCATRYTGLLDALADVADRGEEAAAAD-ELVAF-----LPMKLL-----178
DB 132 -----MEAFARQAHSDVEYSGAEDTHVAVRMSAIALTTGLVLLTLA 174
OY 179 --TFDQOIAMDTYIDS--YAORLHD-EIDSRQELANAVATHEAPLSLEATSDQVABRT 233
DB 175 GWLEVRRAVLRPLERAGHHFDRIADGDLTARLEVRS--ANEIGALFPAALKRMQELTTRI 232
OY 224 DTMRAATDDOYDRMADVSRREISSVASVEVASTADVARTSEDAEALAOQ--EAAAD 291
DB 223 AVMRGVDEINVGAAEISAGNANLSSRTEQAAALEETAAETMEELATTVKQADNAQAQAN 292
OY 292 ALA-----TMDIDEAITGCVTACVEQLGERAADVESYTCVIDDIAEQTNMLAL 339
DB 293 QLAASMOVAORGEVSVAQVQTMHGISASSRQIAD-----IVTVIDGIAFQTNIIAL 345
OY 340 NASIEARAGEAGEGFVAVADEVKALAESRQSTREVELVEQMOAETRETVDOLEVNO 399
DB 346 NAAVMAAAGAEQKGFAYVAGEVRSLAORAAQAEIKALITSSVAITYRAGSQOVAASAG 405
OY 400 RIGEGVEVEEAMETLOETITDAVEDAASQMOEVSTA-----TDEQAVSTEVEAEMVDG 452
DB 406 TMDEVVASVQVRADIMGELISAAQASGIDQVSLAISQMDETTQONALVYQAAAATA 465
OY 453 VDDRAGEITAA 463
DB 466 MEEQARHLAA 476

RESULT 10
ABG24223
ID ABG24223 standard; Protein: 1137 AA.
AC ABG24223;
DT 18-FEB-2002 (first entry)
DE Novel human diagnostic protein #24214.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.
OS
PN WO200175067-A2.
PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX DR N-PSDB; AAS88410.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity
```

```
XX
PS Claim 20; SEQ ID NO 54582; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG3037 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 1137 AA;
Query Match 12.4%; Score 297 5; DB 22; Length 1137;
Best Local Similarity 29.0%; Pred. No. 2.6e-13;
Matches 85; Conservative 61; Mismatches 110; Indels 37; Gaps 7;

OY 204 ROELANAVATHEAPLSLEATSDQVABRTDTR-----ARTDOYDRMADVSRREISS 256
DB 30 RNEI-TAIFASIKTMQALRGTVSDVRKGSQMGHIGIEIYAGNNDLSRTEQQAASLQ 88
OY 257 VSASVEEYAST---ADDVRTSEDAEALAOGEFAADALATMTDIDEAITGVTAGVEQ 312
DB 89 TAASMEQLATVYQONADNRQASSELAKKNATAAQAAGVQ-VSTWHT-----MOE 137
OY 313 LGERAADVESYTCVIDDIAEQTNMLALNASTARAGEGSEFVAVADEVKALAESRQ 372
DB 138 IATSSQKIGDIIISYIDGIAFQTNIIALNAAVEAARAGEGRGFAYVAGEVRMLASRQA 197
OY 373 SRRVEELVEQMOAETRETVDOLEVNORIGEGVEVEEAMETLOETITDAVEDAASQMOEV 432
DB 198 AKEIKGLIE-----ESVN-RVQGSKLVNNAATMIDIVSSVTVDINDIMGEI 243
OY 433 STATDEQAVSTEVEAEMVDGVDDRAGEITAAALDIDADATDQVRTVEEYRETV 485
DB 244 ASASEQQRGIEQVAVQANVSQMDQVYQONASLVEEAAVATEQLANGADHLSSRV 296

RESULT 11
AAM98798
ID AAM98798 standard; Protein: 565 AA.
AC AAM98798;
DT 31-MAR-1999 (first entry)
DE H. pylori GHP0 1282 protein.
XX
XX GHP0 protein; Helicobacter infection; gastroduodenal disease; gastritis;
XX peptic ulcer disease.
XX Helicobacter pylori.
XX
XX MO9843478-A1.
XX
XX 08-OCT-1998.
XX PD
XX 01-APR-1998; 98WO-US06371.
XX
```


Best Local Similarity 23.0%; Pred. No. 1.2e-12;
Matches 101; Conservative 94; Mismatches 145; Indels 100; Gaps 14;

```

OY 70 DHESEYRTODLFANSTKYEOLEKTOAE-----YLLGLGGEVDTEYAARARIGKIHVY 125
XX 70 DHESEYRTODLFANSTKYEOLEKTOAE-----YLLGLGGEVDTEYAARARIGKIHVY 125
XX 70 DHESEYRTODLFANSTKYEOLEKTOAE-----YLLGLGGEVDTEYAARARIGKIHVY 125
DB 206 EYLDPFSSHKENFLA--VEYFKMLGKTESKDNLMMLIAL-TIEKDKYV---EYGVSVRFV 258
OY 126 LGGLPDIYLGAYTRYTGGLDALADVDVADRGEEAAAVDELVARFLPML-----175
DB 259 VIITASAIWALALIATITLMLRA-----IVSSRL-----AVSSTLSHFPLKLNQANSNGIK 310
OY 176 -----KLTTFDQOIANDTY-----IDSYAQR-----196
DB 311 LIEAKSNDELGRMGTAINKNILQTKIMQEDRQAVODTIKIVSDYKAGNAFVITAPAS 370
OY 197 -----LHDEIDSRQE--LANAVATHVEAPLSSLEATS-----QDAERTDPMRATD 242
DB 371 PDLEKELDALNGIMDYLOESVGHMPSIFKIFESYSGLDFRGRIQNASGVELVTNALGQ 430
OY 243 QVDRMADVSRHSSVSASVEEVAADVARTSEDAEALAQGEAAADALATMTDIDEA 302
DB 431 EIQKMLETS-----SNFAKDLANDSANLKECVQNLKASNSQKSLMETSKTIENITTS 484
OY 303 TDGVTAGVEQLGERADVESVTVGVIDIAEQTMALNLSIEARAGEGFAVVADEV 362
DB 485 IQGVSSQSEAMIEQODIKSIVEIRIDIADQTMALNLSIEARAGEHGRGFAVVADEV 544
OY 363 KALAESREGSTRVEE-----LVQOMQAEFEYDQJDEVNQRIGGEVEREEMETLOEI 418
DB 545 RKLAEFRQKLSLEIANINILVOSIDTSISIKNOYKE-----VEETINASIEALRSV 596
OY 419 TDAVEDAASGMQEVSTATDE 438
DB 597 TEGNLKIASDSLEISQEI 616

RESULT 15
AAW71553
ID AAW71553 standard; Protein; 630 AA.
XX
XX AAW71553;
XX
XX 09-NOV-1998 (first entry)
XX
XX Helicobacter polypeptide GHPO 1278.
DE Helicobacter polypeptide GHPO 1278.
XX
XX GHPO 1278; Infection; therapy; diagnosis; vaccine; gastritis;
KM ulcer.
XX
XX Helicobacter pylori.
OS
XX
XX W09821225-A1.
PN
XX
XX 22-MAY-1998.
PD
XX
XX 14-NOV-1997; 97WO-US21353.
PE
XX
XX 29-JUL-1997; 97US-0902615.
PR 14-NOV-1996; 96US-0749051.
PR 01-APR-1997; 97US-0831309.
PR 01-APR-1997; 97US-0833457.
PR 01-APR-1997; 97US-0834705.
PR 24-JUN-1997; 97US-0881227.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA (INNR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
XX
XX Al-Garawi A, Haas R, Kleanthous H, Meyer T, Miller C;
PI Odenbreit S, Tomb J;
XX
XX WPI: 1998-297855/26.
DR N-PSDB; AAV52088.

```

```

XX Helicobacter polynucleotide and polypeptide sequences - useful to
PT treat or prevent gastrointestinal infection
PS Claim 1; Page 308-310; 362pp; English.
XX
XX This claimed Helicobacter pylori polypeptide, designated GHPO 1278,
CC can be used in vaccination methods for preventing or treating
CC Helicobacter infection. 85 Helicobacter polypeptides (see
CC AAW71474-W71558) are claimed, as well as isolated polynucleotides
CC (see AAV52009-93) that encode them. The invention also provides:
CC methods for producing these Helicobacter polypeptides in
CC recombinant host systems, and related expression cassettes, vectors
CC and transformed or transfected host cells; live vaccine vectors
CC that contain the polynucleotides of the invention and which can be
CC used to prevent or treat Helicobacter infection; therapeutic and/or
CC prophylactic methods involving administration of polynucleotide
CC molecules, polypeptides or monospecific antibodies; methods for
CC detecting the presence of Helicobacter in samples using e.g.
CC the polypeptides or monospecific antibodies; and methods for
CC purifying the polypeptides by antibody-based affinity
CC chromatography.
XX
SQ Sequence 630 AA:
Query Match 11.9%; Score 284; DB 19; Length 630;
Best Local Similarity 23.0%; Pred. No. 1.2e-12;
Matches 101; Conservative 94; Mismatches 145; Indels 100; Gaps 14;

OY 70 DHESEYRTODLFANSTKYEOLEKTOAE-----YLLGLGGEVDTEYAARARIGKIHVY 125
DB 206 EYLDPFSSHKENFLA--VEYFKMLGKTESKDNLMMLIAL-TIEKDKYV---EYGVSVRFV 258
OY 126 LGGLPDIYLGAYTRYTGGLDALADVDVADRGEEAAAVDELVARFLPML-----175
DB 259 VIITASAIWALALIATITLMLRA-----IVSSRL-----AVSSTLSHFPLKLNQANSNGIK 310
OY 176 -----KLTTFDQOIANDTY-----IDSYAQR-----196
DB 311 LIEAKSNDELGRMGTAINKNILQTKIMQEDRQAVODTIKIVSDYKAGNAFVITAPAS 370
OY 197 -----LHDEIDSRQE--LANAVATHVEAPLSSLEATS-----QDAERTDPMRATD 242
DB 371 PDLEKELDALNGIMDYLOESVGHMPSIFKIFESYSGLDFRGRIQNASGVELVTNALGQ 430
OY 243 QVDRMADVSRHSSVSASVEEVAADVARTSEDAEALAQGEAAADALATMTDIDEA 302
DB 431 EIQKMLETS-----SNFAKDLANDSANLKECVQNLKASNSQKSLMETSKTIENITTS 484
OY 303 TDGVTAGVEQLGERADVESVTVGVIDIAEQTMALNLSIEARAGEGFAVVADEV 362
DB 485 IQGVSSQSEAMIEQODIKSIVEIRIDIADQTMALNLSIEARAGEHGRGFAVVADEV 544
OY 363 KALAESREGSTRVEE-----LVQOMQAEFEYDQJDEVNQRIGGEVEREEMETLOEI 418
DB 545 RKLAEFRQKLSLEIANINILVOSIDTSISIKNOYKE-----VEETINASIEALRSV 596
OY 419 TDAVEDAASGMQEVSTATDE 438
DB 597 TEGNLKIASDSLEISQEI 616

```

Search completed: January 2, 2003, 12:41:39
Job time : 99 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2003, 12:29:51 ; Search time 17 Seconds
(without alignments)
846.342 Million cell updates/sec

Title: US-09-455-978B-2
Perfect score: 2394
Sequence: 1 MSNDNDTLVADVRNGIDGH.....ATDQVRVEEVRVTKLS 489

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTUTS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	205	8.6	955	2 US-08-428-414A-3	Sequence 3, Appl1
2	202.5	8.5	955	1 US-08-006-676B-1	Sequence 1, Appl1
3	202.5	8.5	955	1 US-08-282-845-2	Sequence 2, Appl1
4	202.5	8.5	955	5 PCT-US94-00324-1	Sequence 1, Appl1
5	194	8.1	1786	4 US-08-973-462-8	Sequence 8, Appl1
6	188.5	7.9	885	2 US-08-533-306A-4	Sequence 4, Appl1
7	188.5	7.9	885	2 US-08-742-923A-4	Sequence 4, Appl1
8	187	7.8	630	4 US-08-973-462-9	Sequence 4, Appl1
9	178	7.4	2101	1 US-08-466-390-4	Sequence 4, Appl1
10	178	7.4	2101	1 US-08-470-950-4	Sequence 4, Appl1
11	178	7.4	2101	1 US-08-467-781-4	Sequence 4, Appl1
12	178	7.4	2101	1 US-08-195-487-4	Sequence 4, Appl1
13	178	7.4	2101	2 US-08-483-924-4	Sequence 4, Appl1
14	178	7.4	2101	4 US-09-452-294-1	Sequence 1, Appl1
15	178	7.4	2101	5 PCT-US93-06160-4	Sequence 4, Appl1
16	176	7.4	540	4 US-08-973-462-22	Sequence 22, Appl1
17	173	7.2	1196	1 US-08-144-121-4	Sequence 4, Appl1
18	173	7.2	1196	2 US-08-735-883-4	Sequence 4, Appl1
19	170	7.1	816	2 US-08-533-306A-6	Sequence 6, Appl1
20	170	7.1	816	2 US-08-742-923A-6	Sequence 6, Appl1
21	164.5	6.9	534	4 US-09-103-664A-2	Sequence 2, Appl1
22	160.5	6.7	1939	4 US-09-310-187A-1	Sequence 1, Appl1
23	159	6.6	1018	1 US-08-072-610-2	Sequence 2, Appl1
24	159	6.6	1018	2 US-08-719-892B-2	Sequence 2, Appl1
25	159	6.6	1018	4 US-09-092-458-2	Sequence 2, Appl1
26	156	6.5	1886	4 US-08-938-105-3	Sequence 3, Appl1
27	156	6.5	3696	4 US-09-134-001C-5080	Sequence 5080, Ap

28	154.5	6.5	1354	3 US-08-685-871-2	Sequence 2, Appl1
29	152	6.3	576	2 US-08-533-306A-2	Sequence 2, Appl1
30	152	6.3	576	2 US-08-742-923A-2	Sequence 2, Appl1
31	151	6.3	1184	4 US-09-541-782-2	Sequence 2, Appl1
32	151	6.3	1184	4 US-09-723-860-2	Sequence 2, Appl1
33	151	6.3	2482	1 US-08-328-254-6	Sequence 6, Appl1
34	151	6.3	3111	2 US-08-460-309-4	Sequence 4, Appl1
35	151	6.3	3111	2 US-08-125-077-4	Sequence 4, Appl1
36	150.5	6.3	566	2 US-08-533-669A-8	Sequence 8, Appl1
37	150.5	6.3	566	2 US-08-511-872-2	Sequence 8, Appl1
38	150.5	6.3	566	4 US-09-183-861-8	Sequence 8, Appl1
39	150.5	6.3	566	4 US-09-022-765-8	Sequence 8, Appl1
40	150	6.3	527	2 US-08-762-106-9	Sequence 9, Appl1
41	150	6.3	527	4 US-09-320-774-9	Sequence 9, Appl1
42	150	6.3	941	4 US-09-336-447A-9	Sequence 9, Appl1
43	150	6.3	3248	1 US-08-353-700-1	Sequence 1, Appl1
44	150	6.3	3248	5 PCT-US95-16216-1	Sequence 1, Appl1
45	149	6.2	516	2 US-08-762-106-8	Sequence 8, Appl1

ALIGNMENTS

RESULT 1
US-08-428-414A-3
Sequence 3, Application US/08428414A
Patent No. 5912166
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
TITLE OF INVENTION: LEISHMANIASIS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESS: SEED AND BERRY
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/428,414A
APPLICATION NUMBER: US/08/428,414A
FILING DATE: 21-Apr-1995
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Radlecek, Ann T.
REGISTRATION NUMBER: 39,244
REFERENCE/DOCKET NUMBER: 210121.407
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 682-4900
TELEFAX: (206) 682-6031
TELEX: 3723836 SEEDANDBERRY
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 955 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-428-414A-3
Query Match 8.6%, Score 205; DB 2; Length 955;
Best Local Similarity 23.3%; Pred. No. 6e-08;
Matches 115; Conservative 80; Mismatches 198; Indels 100; Gaps 19;
QY 13 VANGIDGHALDRICLDEEIMRSLFSGIDDTJMAALAEQPLREATDAVT---DFY 69
DB 530 VARRIDAETASERKL-----ESTVAQLEREBQREREVALDALQTHQRKLO 574
QY 70 DILSEYERTQDUFANSTFTVEQLKETQAEYLLGLGRGEYDREYAAQRRARICKIHDLVLGIG 129

```

Db      575 EALSESSESTA--AEROLLQQLLEOSE-----RTQL-SQVYTBRERLTR--DLQRT- 621
Qy      130 PDVYLGAATRRYTGLLDALDDVAVADRGEERAAAADVLVARFLPMKLITLFDOQIAMDY 189
Db      622 -----QGEYGETE-LARDVALCAAGQEWARYHAAVPHLQTLLELAT-----EW 663
Qy      190 IDSYAQRILHDIIDRSQELANNAVATHVAPLSSLEATISQDVAERPDYTRATDDQVBRAD 249
Db      664 EDALRERLARLD-----EAAAELLDAASTSQNARESACERTLSLEQLRSEERAAE 717
Qy      250 VSREISSVSASVEEAVSTADVDRTSD-----AEALAAQGEAADDALATPMTDIDE 301
Db      718 LASQLEATIAAAKSSAEQDRETRATLEQQLRESEARAELASQLEATIAAAKSSAEQDREN 777
Qy      302 ATDGVTAGVEQLGRRADVESVTGVIDIAEQTMILAN--SIEARAGEAGEGFAYVA 359
Db      778 TRATLEQQLRBSERRAAELASQ-----LESTIAAKMSAEQDRESTR----- 819
Qy      360 DEYKALAESEESQSTREVELVEQMAET--EETVDQIDE-----VNORIGEGYERVEEAM 412
Db      820 -----TLEQQLRDSERRAAELASQLESTIAAKMSAEQDRESTRATLEQQLRESEARAELA 875
Qy      413 ETLOEITDAVDAASGMOEVSTATDEAV--STEEVAVENQYGVDDRAGETIAALLDIAD- 469
Db      876 SOL-ESTIAAKMSAEQDRESTRATLEQQLRDSERRAAELASQLEATIAAAKSSAEQDREN 934
Qy      470 --ATDQVRYVEE 480
Db      935 RAALQQLRDSRE 947

```

RESULTS
US-08-006-676B-1
; Sequence 1, Application US/08006676B

```

: GENERAL INFORMATION:
:
: APPLICANT: Reed, Steven.
:
: TITLE OF INVENTION: Diagnosis of Leishmaniasis
:
: NUMBER OF SEQUENCES: 3
:
: CORRESPONDENCE ADDRESS:
:

```

ADDRESSEE: Jeffrey B. Oster
STREET: 8339 SE 57th Street
CITY: Mercer Island
STATE: Washington

```

; COUNTRY: USA
; ZIP: 98040-4906
; COMPUTER READABLE FORM:

```

```

: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: WORD for Windows
: CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/006,676B
FILING DATE: 15-JAN-1993

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: Oster, Jeffrey B.
REGISTRATION NUMBER: 32,585

REFERENCE/DOCKET NUMBER: REF ID: A66062
TELECOMMUNICATION INFORMATION: TELEPHONE: 408-222-7045

TELEPHONE: (206) 232 7843
TELEFAX: (206) 236 0205
INFORMATION FOR CEO: TD NO. 1

```

; INFORMATION FOR SEQ ID NO: 1:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 955 amino acids

```

LENGTH: 500 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-006-676B-1

Query Match
Best Local Similarity

```
8.5%; Score 202.5; DB 1; Length 955;  
23.8%; Pred. No. 9.4e-08;
```

	Matches	110;	Conservative	77;	Mismatches	190;	Indels	85;	Gaps	18;
Qy	44	DDTMAALAAEQFLFATADALVT---	DDYDHESEYERODLFANSTKYVEOLKETQAEYL	100						
Db	546	ESTVAQLREQGERVALDALQTHORKLOEALSESRTA--	AERDQLQLQTEILOSE--	600						
Qy	101	LGIGGEVDTEAARARIGKIHVDLGLGDPVYLCAVRYTYGLDALADVDVADGEEA	160							
Db	601	---RTQL-SQVYVDREKTR--DLQRI-----	QYXGGENE-LARDVALCAQGM	642						
Qy	161	AAAVDELARFLPMKLTLFDQQLAMDYITDSYAQRHLDEIDSRQLANAVATYHEAPLS	220							
Db	643	EARYHAAFVHLQTLLELAT-----	EEDDLRRALLAERQ-----	EAAAEELDAANS	688					
Qy	221	SLEATSGOVAERTDPMRARTIDQVRMADVSRREISSVASVEEVASTADVDRTSED---	277							
Db	669	TSQNNRESACERLTLEEQLRSEERRAEALASQLLENTAAKSSAQDRENTATYLEQQLR	748							
Qy	278	----AEALAQGEAAADALATMTDIDEANTDGTAGYVEQLGERADAVESVTGVIDDIAE	332							
Db	749	ESEARAARELASQLLENTAAKMSAQDRENTATYLEQQLDSEERRAEALASQ-----	LE	801						
Qy	333	QTNMLALNA--SIEAARGENGEGRAVVADEVKALAEESREGSTVEELVEGMOAET---E	388							
Db	802	STTAAKMSAQDRESTR-----	TLQQLRQSEERRAEALASQLESTTAAK	846						
Qy	389	ETVDQDLDE-----VNRIGEGVEERYEAMETLOETITDAVEDAASGQMEVSTATDQAV--	441							
Db	847	MSAQDRESTRATYLEQQLRSEERRAEALASQL-ESTTAAKMSAQDRESTRATYLEQQLND	905							
Qy	442	STEEVAENVADVDDRAGETIAALDITAD---ATDQOVRVEE	480							
Db	906	SEERAARELASQLENTAAKSSAQDRENTRALLEQQLRSE	947							

RESULT 3
US-08-282-845-2

; Sequence 2, Application US/08282845
; Patent NO. 5719263
; CURRENT INVENTION

;; GENERAL INFORMATION:
;;
; APPLICANT: Reed, Steven G
; TITLE OF INVENTION: A 3200

1 TITLE OF INVENTION: A 250kd antigen present in Leishmania
2
3 TITLE OF INVENTION: Species
4
5 TITLE OF INVENTION: Species
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834

NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESS: Immunex Corporation

ADDRESS: Insurance Corporation
STREET: 51 University Street
CITY: Seattle

STATE: WA
COUNTRY: USA

ZIP: 98101
COMPUTER READABLE

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh

```

```
;; OPERATING SYSTEM: Apple Macintosh Operating System 7.1
;; SOFTWARE: Microsoft Word for Macintosh 5.1a
```

```

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/282,845
;

```

```

; FILING DATE:
; CLASSIFICATION: 435
;

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/006,676
;

```

FILING DATE: JANUARY 1971
CLASSIFICATION: 435
; ;
; ;
; ;

ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34 603

REGISTRATION NUMBER: 34,053
REFERENCE NUMBER: 5004-P
TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644

```

; INFORMATION FOR SEQ ID NO: 2:
;
; SEQUENCE CHARACTERISTICS:

```

Table 1

LENGTH: 955 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-282-845-2

Query Match 8.5%; Score 202.5; DB 1; Length 955;
Best Local Similarity 23.8%; Pred. No. 9.4e-08;
Matches 110; Conservative 77; Mismatches 190; Indels 85; Gaps 18;

QY 44 DDTMALAEOPLFEATADALVT---DEYDHESTERTODLFANSTKVEQLKETQAEYL 100
DB 546 ESTVAQLERQERREVALDALQTHQRKLEALESESRTA---AERDQLLOQLTEIQSE-- 600
QY 101 LGIGGEVDTEYAQRARIGKIHVDVGLGPDVYLCATRYTGLDALADVDVADRGEA 160
DB 601 ---RTQL-SQVYDRERLTR--DLQRI-----QYEGETE-LARDVALCAQDEM 642
QY 161 AAADVELARFLPMLKLTFFDQIAMDYIDSYAQRHDEIDSRQELANAVATHEAPLS 220
DB 643 EARYNAVFHLOTLELAT-----EMEDALRERALAERD-----EAAAELDMAAS 688
QY 221 SLEATSQVAERTDMRATDQVDRMADVSRHSSVSASVEAVSTADVVRISD--- 277
DB 689 TSQNRASACERLTSLQQLRESEERAAELASQLBATAAKSSAQDRENTRATLEQQLR 748
QY 278 -----AEALAOGEAAADALATMTDIDEATDGVAGVQLGERADVESYGVDDIAE 332
DB 749 ESEARAELASQLEATTAAKMSAQEDRENTRATLEQQLRSEERAAELASQ-----LE 801
QY 333 QTNMLALNA--SIEARAGEGEPFAVVADEVKALAESRQSTVEELVEQMAET--E 388
DB 802 STTAKMSAQEDRESTR-----TEEQRLRSEERAAELASQLESTTAK 846
QY 389 ETVDQLDE---VNQIGEGEVEREAMETLQETITDAVEDASQMOEVSTATDQAV-- 441
DB 847 MSAEQDRESTRATLEQQLRESEERAAELASQL-ESTTAKMSAQEDRESTRATLEQQLRD 905
QY 442 STEEVAENVDGVDRAAGETAAALDDIAD---ATDQVTRVEE 480
DB 906 SEERAELASQLEATTAAKSSAQEDRENTRATLEQQLRSE 947

RESULT 4

PCT-US94-00324-1
Sequence 1, Application PC/TUS9400324
GENERAL INFORMATION:
APPLICANT: Reed, Steven
TITLE OF INVENTION: Diagnosis of Leishmaniasis
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: Apple Macintosh
SOFTWARE: Microsoft Word, version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/00324
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/006, 676
FILING DATE: 15-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 5004-WO

TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 955 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-00324-1

Query Match 8.5%; Score 202.5; DB 5; Length 955;
Best Local Similarity 23.8%; Pred. No. 9.4e-08;
Matches 110; Conservative 77; Mismatches 190; Indels 85; Gaps 18;

QY 44 DDTMALAEOPLFEATADALVT---DEYDHESTERTODLFANSTKVEQLKETQAEYL 100
DB 546 ESTVAQLERQERREVALDALQTHQRKLEALESESRTA---AERDQLLOQLTEIQSE-- 600
QY 101 LGIGGEVDTEYAQRARIGKIHVDVGLGPDVYLCATRYTGLDALADVDVADRGEA 160
DB 601 ---RTQL-SQVYDRERLTR--DLQRI-----QYEGETE-LARDVALCAQDEM 642
QY 161 AAADVELARFLPMLKLTFFDQIAMDYIDSYAQRHDEIDSRQELANAVATHEAPLS 220
DB 643 EARYNAVFHLOTLELAT-----EMEDALRERALAERD-----EAAAELDMAAS 688
QY 221 SLEATSQVAERTDMRATDQVDRMADVSRHSSVSASVEAVSTADVVRISD--- 277
DB 689 TSQNRASACERLTSLQQLRESEERAAELASQLBATAAKSSAQDRENTRATLEQQLR 748
QY 278 -----AEALAOGEAAADALATMTDIDEATDGVAGVQLGERADVESYGVDDIAE 332
DB 749 ESEARAELASQLEATTAAKMSAQEDRENTRATLEQQLRSEERAAELASQ-----LE 801
QY 333 QTNMLALNA--SIEARAGEGEPFAVVADEVKALAESRQSTVEELVEQMAET--E 388
DB 802 STTAKMSAQEDRESTR-----TEEQRLRSEERAAELASQLESTTAK 846
QY 389 ETVDQLDE---VNQIGEGEVEREAMETLQETITDAVEDASQMOEVSTATDQAV-- 441
DB 847 MSAEQDRESTRATLEQQLRESEERAAELASQL-ESTTAKMSAQEDRESTRATLEQQLRD 905
QY 442 STEEVAENVDGVDRAAGETAAALDDIAD---ATDQVTRVEE 480
DB 906 SEERAELASQLEATTAAKSSAQEDRENTRATLEQQLRSE 947

RESULT 5

US-08-973-462-8
Sequence 8, Application US/089734628
Patent No. 6191270
GENERAL INFORMATION:
APPLICANT: DRUILHE, PIERRE
TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
FILE REFERENCE: 0660-0125-0 PCT
CURRENT APPLICATION NUMBER: US/08/973, 462B
CURRENT FILING DATE: 1998-02-06
EARLIER APPLICATION NUMBER: PCT/FR96/00894
EARLIER FILING DATE: 1996-06-12
EARLIER APPLICATION NUMBER: FR 95/07007
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 1786
TYPE: PPT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Polypeptide
US-08-973-462-8

NAME: Smith, Deann F.
REGISTRATION NUMBER: 36683
REFERENCE/DOCKET NUMBER: 2115-00869DVC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 641-1500
TELEFAX: (810) 641-0270
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 885 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-742-923a-4

Query Match 7.9% Score 188.5; DB 2; Length 885;
Best Local Similarity 21.7%; Pred. No. 1.1e-06;
Matches 113; Conservative 98; Mismatches 178; Indels 131; Gaps 23;

QY 44 DDTMAALAEQPLFEATDALVTDFYDHLSEYERTODLFANSTKTVEOLKETOAEY---L 100
DB 134 EDALAAQAFEE-----ARRTRFEEDRDRSHREMEKAKNDKKOTLEKNMDLAEEL 187
QY 101 LGLGRGYDTERYAARARIGIHVYLGIPDYVYLGAVTRYTGILLDALADVVA--DRGE 158
DB 188 RVLGQAKQEVHEKKKK-----LEAQVQLQSKSDGE 219
QY 159 EAAAVAVDELVARFLPMKLLT-----PDQQLAMDTYIDSYAQRLLD-----EIDSRQL 207
DB 220 RARLELNKVKHKLQNEVSIVGMLNEAGKAIKLAKDVASLSQQLDTQELLQEEETROKL 279
QY 208 ANAVATHEAPLSLEASQDAERTDTMRARTDDQVDMADVSRISVS-----AS 260
DB 280 -----NSTKLRLQEL-----ERNSLQQLDDEEMAKQNLNHTITLNLQSLDSKK 326
QY 261 VEEVASTADV---RTSEDAEALAOGE--AAADAL-ATMTDIDBATDGTAGVAGQL 313
DB 327 LQDFASTVEALEEGKRRFQKEITENLTQYEKAAAYDKLETKTKNLQQLDLDLVLDNQ 386
QY 314 GERADVSVTGVIDD-TAEQTNMLALNASTEAARAGAGEGFANVADEVKALA-----E 367
DB 387 RQVYSNLEKKRRFPQDLAEEKNISKYADERDRAEAARE-----KETALSLARALE 440
QY 368 ESREOSTREVELVEQMAETEETVDQLDENVQRIGEGYERVEAMET-----LOBET 419
DB 441 EALEKKELEETKMKLAKMEDIVSKDVGKNVHE-LEKSRALFETQMEEMKKTLELE 499
QY 420 D---AVEADA---ASGQGEVS-----TATDEQAVST-----BEVAMVDGVDR 456
DB 500 DELQASEDAKLRLLEVNMQALKGQFERDQLQARDQONEKRRQLQRLHVEYTELEDERNER 559
QY 457 AGEITAA-----LDDIADATDQOVRTVEEVRTVGKL 488
DB 560 A--LAAAKKKLEGLDKLELQADSATKGREBATKQLRKL 597

RESULT 8
US-08-973-462-9
Sequence 9, Application US/08973462B
Patent No. 6191270
GENERAL INFORMATION:
APPLICANT: DRUILLHE, PIERRE
APPLICANT: DAUBERSIES, PIERRE
TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
FILE REFERENCE: 0660-0123-0 PCT
CURRENT APPLICATION NUMBER: US/08/973.462B
CURRENT FILING DATE: 1998-02-06
EARLIER APPLICATION NUMBER: PCT/FR96/00894
EARLIER FILING DATE: 1996-06-12
EARLIER APPLICATION NUMBER: FR 95/07007
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn ver. 2.0
SEQ ID NO 9

LENGTH: 630
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Polypeptide
US-08-973-462-9

Query Match 7.8% Score 187 DB 4; Length 630;
Best Local Similarity 21.5%; Pred. No. 8.8e-07;
Matches 85; Conservative 87; Mismatches 138; Indels 86; Gaps 16;

QY 150 DDVVAADRGEEAAA-AVDELVARFLPMK-LLTTPDQQLAMDTYIDSYAQRLLDDEIDSRQL 206
DB 263 EEIVAPVEESVAPVEEIVA---PTVESVAPVEEIVPSVESVAPVEESV--AEN 317
QY 207 LANAVATHEAPLSLETSODVAERTDTMRARTDDQVDMADVSRISVSASVSEVVA- 265
DB 318 VEESVAVNEVEE--SVAENVEESVAVNEESVAVNEE-----VAPSVEEIVA 363
QY 266 -----STADVVRTSEDAEALAO-----QGEAAADALATMTDIDEATDGTAGVAGQLGE 315
DB 364 PTVESVAVENVAATNLSD-NLSNLLGIGIETEBIIDSILNEIEVKE--NVTTTIEKVEE 420
QY 316 RAAD-VESVTGVIDDIABOT-----NMLALNASTEAARAGAGEGFANVADE 361
DB 421 TTAEVTFEFSNLEIEQENTITNDTIEEKLDELHENVLSALENTQSEKKEVIDYIEE 480
QY 362 VKALAESREOSTREVELVEQMAETEETVDQ-----DEVNQRIGEGYERVEAM- 412
DB 481 VK-----EEVATVLTIEVEQAESEESTITEIFENLEENAVESNEKVAENLEKLETVF 534
QY 413 -ETLQETIDAVE-----DAASGQEVSTADDEQAVSTEEVAAEAVD 451
DB 535 NTVIDKVEETVEIGESGLENNEMDKAFSEIFDNVKGIOENLLGMRFSIETSTVIOSEE 594
QY 452 GVDNRAGEIAAALDIADATDQOVRTVEEVRTVGK 487
DB 595 KVDLNEVYSSLDNIENMKGLKLEKNISSTSEE 630

RESULT 9
US-08-466-390-4
Sequence 4, Application US/08466390
Patent No. 5686562
GENERAL INFORMATION:
APPLICANT: TOKURITLY, GARY
APPLICANT: LIDGARD, GRAHAM P
TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HORWITZ & THIBEAULT
STREET: 125 HIGH STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466.390
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER ESO, EDMUND R
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: MTP-013
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100


```

; APPLICATION NUMBER: US/08/467,781
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER ESQ, EDMUND R
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: MTP-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2101 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-467-781-4

Query Match          7.4%; Score 178; DB 1; Length 2101;
Best Local Similarity 18.9%; Pred. No. 2.2e-05;
Matches 86; Conservative 95; Mismatches 192; Indels 82; Gaps 14;

QY 67 DRYDHLSE-YERTODLFANSTKTVBOLKETOAEYLLGLGRGEYDTEYAAQRARI--GKIH 123
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 329 EFASHLQOLQDALNELTEHRSKATQEMLEKQALELSAALQDKCKLEKEKNEILLQKLS 388
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 124 -----DVLGLGPDVYLGATRYTGTLLDALADVDVADRGEEAAAVD 165
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 389 QLEEHLSQLQDNPPQEKGEVLG-----DVLQLETLKQEAATLAAANT 430
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 166 ELVARFLPLMKLLTFDQQLAMDTYIDSYAORLHDEIDSROELANAVATVEAPLSLEAT 225
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 431 QLOAR-VEMLETERGQGEKLT-----LAERGHFE-EKKQQLSSLI-TDLOSSISLSQA 481
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 226 SODVAERTDMARTDDOVDRMADYSREISSVASYEEVASTADVRRTSSEDAEALAOQG 285
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 482 KEELQASQAHGARLTAQV--ASLTSELTTLNATIQQ-----ODQLAGLKQQA 528
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 286 EAAADALATMTDIDATDGTAGVEQLGERADVESYTGVIDIAEQTNMLALNASIEA 345
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 529 KKKQALQTLQOOEQASGLRHQYEQSLSSLKQEQ--QLKEVAEKQEAATRDHQAOL 585
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 346 ARAGEGEGFAVVAD-EVKALAEESREOSTRYEELVEQQAETE-----ETVDQLDEVN 398
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 586 ATAABERESLSRERDAALKQLEALEKEKAKLEILLQOQLQVANEARDSAQTSTVQAQRK 645
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 399 QRIEGGVERVEEMETLQETIDAVEDDASGMOEVSATDEQAVSTEEVAEMVDGVDDR-- 456
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 646 AELSKRVEELQCVETARQEQHEAQOVALELEQLRSEQQKATEKERVAAQEKDQLQEQLO 705
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 457 -----AGEIAAALDDIADATDQOVRTVEVR 482
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 706 ALKESLKVTKGSLSEKRRRAADALEEQRCISELK 740
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
US-08-195-487-4
; Sequence 4, Application US/08195487
; Patent No. 5783403
; GENERAL INFORMATION:
; APPLICANT: TOURKATLY, GARY
; APPLICANT: LIDGARD, GRAHAM P
; TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
; TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA HURWITZ & THIBEAULT
; STREET: 53 STATE STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/195,487
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/901,701
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER ESQ, EDMUND R
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: MTP-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/248-7100
; TELEFAX: 617/248-7000
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2101 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-195-487-4

Query Match          7.4%; Score 178; DB 1; Length 2101;
Best Local Similarity 18.9%; Pred. No. 2.2e-05;
Matches 86; Conservative 95; Mismatches 192; Indels 82; Gaps 14;

QY 67 DRYDHLSE-YERTODLFANSTKTVBOLKETOAEYLLGLGRGEYDTEYAAQRARI--GKIH 123
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 329 EFASHLQOLQDALNELTEHRSKATQEMLEKQALELSAALQDKCKLEKEKNEILLQKLS 388
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 124 -----DVLGLGPDVYLGATRYTGTLLDALADVDVADRGEEAAAVD 165
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 389 QLEEHLSQLQDNPPQEKGEVLG-----DVLQLETLKQEAATLAAANT 430
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 166 ELVARFLPLMKLLTFDQQLAMDTYIDSYAORLHDEIDSROELANAVATVEAPLSLEAT 225
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 431 QLOAR-VEMLETERGQGEKLT-----LAERGHFE-EKKQQLSSLI-TDLOSSISLSQA 481
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 226 SODVAERTDMARTDDOVDRMADYSREISSVASYEEVASTADVRRTSSEDAEALAOQG 285
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 482 KEELQASQAHGARLTAQV--ASLTSELTTLNATIQQ-----ODQLAGLKQQA 528
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 286 EAAADALATMTDIDATDGTAGVEQLGERADVESYTGVIDIAEQTNMLALNASIEA 345
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 529 KKKQALQTLQOOEQASGLRHQYEQSLSSLKQEQ--QLKEVAEKQEAATRDHQAOL 585
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 346 ARAGEGEGFAVVAD-EVKALAEESREOSTRYEELVEQQAETE-----ETVDQLDEVN 398
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 586 ATAABERESLSRERDAALKQLEALEKEKAKLEILLQOQLQVANEARDSAQTSTVQAQRK 645
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 399 QRIEGGVERVEEMETLQETIDAVEDDASGMOEVSATDEQAVSTEEVAEMVDGVDDR-- 456
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 646 AELSKRVEELQCVETARQEQHEAQOVALELEQLRSEQQKATEKERVAAQEKDQLQEQLO 705
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 457 -----AGEIAAALDDIADATDQOVRTVEVR 482
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 706 ALKESLKVTKGSLSEKRRRAADALEEQRCISELK 740
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
US-08-483-924-4
; Sequence 4, Application US/08483924
; Patent No. 5882876
; GENERAL INFORMATION:
; APPLICANT: TOURKATLY, GARY
; APPLICANT: LIDGARD, GRAHAM P
; TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
; TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX
; NUMBER OF SEQUENCES: 6
```

```

CORRESPONDENCE ADDRESS:
ADDRESS: TESTA, HURWITZ & THIBEAULT
STREET: 125 HIGH STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patdntln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,924
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER, ESO, EDMUND R
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: MTP-013
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2101 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-483-924-4

```

```

Query Match      7.4%; Score 178; DB 2; Length 2101;
Best Local Similarity 18.9%; Pred. No. 2,2e-05;
Matches 86; Conservative 95; Mismatches 192; Indels 82; Gaps 14;

```

```

QY 67 DFYDHLSE-YERQDLEFANSTKTVBOLKETAQAEYLLGLGGEYDTEYAQRARI--GKIH 123
   :|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 329 EFASHLQLODLDALNELTEHRSKATQEWLEKQAOLEKELSAALODKCKLEEKNETLOGKLS 388
QY 124 -----DVLGLGPDVYLGAATRYTGLDALADLVADRGEEAAAVD 165
   :|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 389 QLEHLSQLDQNPPOKEGVLG-----DVQLLETLKQEAATLANNNT 430
QY 166 ELVARELPMKLKLTFFDQIAMDTYIDSYAQRHDEIDSRQELANAVATHEAPLSLEAT 225
   :|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 431 QLOAR-VEMLETERGQGEAKL-----LABRGHFE-BEKQOLSLI-TDLOSSISNLSQA 481
QY 226 SQVAERTDTRMARTDQVVRMADYSREISSVASVEVASTADVARTSEDAEALAQOG 285
   :|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 482 KEELDAQSQAHGARLTAAQV---ASLTSELTTLNATLTIQO-----ODQELAGLKQA 528
QY 286 EAAADDAATMTDIDETAGTAGEVQLGERADVESVTGVIDDIAEQTMLALNASTEA 345
   :|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 529 KEKQAOQAQTLQOQEQASQGLRHQVEQLSSLKQKEQ---QLEKVAEKQATKQDNHAQOL 585
QY 346 ARAGEAGEGFAVAD-EVKALAEBSREQSTRVEELVEQMAETE-----ETVDQULDEVN 398
   :|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 586 ATAAEEEREAALRRBDALKOLEALEKEKAKEAKLEITLQOQLVANEARDSAGTSVQAOREK 645
QY 399 QRIGEGYERVEEAMETLQETTTDAVEDAASGMQEVSTATDQAVSTEVAVMGVQVDR-- 456
   :|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 646 AELSRKVEELQACVETARQEQHEAQAQVALEQLRSEQKATEKEKRVAAQEKDQLOEQLO 705
QY 457 -----AGEITAAALDDIADTDQOVRTVEEVR 482
   :|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 706 ALKESLKVTKGSLPEEKRRADAALDEQRCISELK 740

```

```

RESULT 14
US-09-452-294-1
Sequence 1, Application US/09452294
Patent No. 6287790
GENERAL INFORMATION:

```

```

APPLICANT: Lelievre, Sophie
APPLICANT: Bissell, Mina
TITLE OF INVENTION: UTILIZATION OF NUCLEAR STRUCTURAL PROTEINS FOR TARGETED
TITLE OF INVENTION: THERAPY AND DETECTION OF PROLIFERATIVE AND
TITLE OF INVENTION: DIFFERENTIATION DISORDERS
FILE REFERENCE: IB-1454- Sequence Submittal
Patent No. 6287790
CURRENT APPLICATION NUMBER: US/09/452,294
CURRENT FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: 60/710,420
PRIOR FILING DATE: 1998-11-30
NUMBER OF SEQ ID NOS: 1
SOFTWARE: Patdntln Ver. 2.1
SEQ ID NO 1
LENGTH: 2101
TYPE: PRT
ORGANISM: Homo sapiens
US-09-452-294-1

```

```

Query Match      7.4%; Score 178; DB 4; Length 2101;
Best Local Similarity 18.9%; Pred. No. 2,2e-05;
Matches 86; Conservative 95; Mismatches 192; Indels 82; Gaps 14;

```

```

QY 67 DFYDHLSE-YERQDLEFANSTKTVBOLKETAQAEYLLGLGGEYDTEYAQRARI--GKIH 123
   :|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 329 EFASHLQLODLDALNELTEHRSKATQEWLEKQAOLEKELSAALODKCKLEEKNETLOGKLS 388
QY 124 -----DVLGLGPDVYLGAATRYTGLDALADLVADRGEEAAAVD 165
   :|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 389 QLEHLSQLDQNPPOKEGVLG-----DVQLLETLKQEAATLANNNT 430
QY 166 ELVARELPMKLKLTFFDQIAMDTYIDSYAQRHDEIDSRQELANAVATHEAPLSLEAT 225
   :|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 431 QLOAR-VEMLETERGQGEAKL-----LABRGHFE-BEKQOLSLI-TDLOSSISNLSQA 481
QY 226 SQVAERTDTRMARTDQVVRMADYSREISSVASVEVASTADVARTSEDAEALAQOG 285
   :|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 482 KEELDAQSQAHGARLTAAQV---ASLTSELTTLNATLTIQO-----ODQELAGLKQA 528
QY 286 EAAADDAATMTDIDETAGTAGEVQLGERADVESVTGVIDDIAEQTMLALNASTEA 345
   :|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 529 KEKQAOQAQTLQOQEQASQGLRHQVEQLSSLKQKEQ---QLEKVAEKQATKQDNHAQOL 585
QY 346 ARAGEAGEGFAVAD-EVKALAEBSREQSTRVEELVEQMAETE-----ETVDQULDEVN 398
   :|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 586 ATAAEEEREAALRRBDALKOLEALEKEKAKEAKLEITLQOQLVANEARDSAGTSVQAOREK 645
QY 399 QRIGEGYERVEEAMETLQETTTDAVEDAASGMQEVSTATDQAVSTEVAVMGVQVDR-- 456
   :|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 646 AELSRKVEELQACVETARQEQHEAQAQVALEQLRSEQKATEKEKRVAAQEKDQLOEQLO 705
QY 457 -----AGEITAAALDDIADTDQOVRTVEEVR 482
   :|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 706 ALKESLKVTKGSLPEEKRRADAALDEQRCISELK 740

```

```

RESULT 15
PCT-US93-06160-4
Sequence 4, Application PC/TUS9306160
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSER: TESTA HURWITZ & THIBEAULT
STREET: 53 STATE STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/06160
FILING DATE: 19930621
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER ESO, EDMUND R
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: MTP-013
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/248-7000
TELEFAX: 617/248-7100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2101 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-06160-4

Query Match 7.48; Score 178; DB 5; Length 2101;

Best Local Similarity 18.9%; Pred. No. 2.2e-05;

Matches 86; Conservative 95; Mismatches 192; Indels 82; Gaps 14;

QY 67 DEYDLES-YERTODLFANSKTVQKETOAEVLLGREGYDTEYAQRARI--GKIH 123
DB 329 EFASHLQQLDQALNELTEBHSKATQEWLEKQALKELSAALQDKKCLEKKEILQKLS 388
QY 124 -----DVLGLGPDVYLGATRYTGLLDALADDVAVADGEEAANAVID 165
DB 389 QLEHLSQLQDNPPQKEGEVLS-----DVLQLETLKQEAATLAANNVT 430
QY 166 ELVARFLPMKLITFDQIAMDYIDSYAQRILHDEISRQELANAVATHVEAPLSLEAT 225
DB 431 QLOAR-VEMLETERGQOEAKL-----LAERGHEE-EKQQLSSLI-TDLQSSISLSQA 481
QY 226 SODVAFRTDMARFTDDQVDRADVSREISSASVEEYASTADVRRTSDEAALAQOG 285
DB 482 KEELQASQANGARLTAQV---ASLTSELTTLNATIQO-----ODOELAGLKQQA 528
QY 286 EAADDA LATMTDIDEATDGTAGVGEOLGERADVESVTGVIDIAEQTNMLALNASTEA 345
DB 529 KEKQALQLOTLOQOEQAQOGRLHVEQLSSLSLKQKEQ---QLKEVAEKQEAATRODHAQOL 585
QY 346 ARAGAGEGFAVVAD-BYKALAEBSREOSTRYEELVEQMAETE-----ETVDQLDEVN 398
DB 586 ATAAERREASLRERDAALKOLEAKLEILQOOLQVANEARDSAQTSTQAQREK 645
QY 399 ORIGGEVREVEAMETLEITDAVEDAASGMQEVSTATDEQAVSTEEVAEMVDGVDDR-- 456
DB 646 AELSRRVLELQACVETARQOEHAQOVAELQLRSEQKATEREVAQEKQOLQOEQLQ 705
QY 457 -----AGEIAAALDDIADATDOQVTVEEVR 482
DB 706 ALKESLKYTKGSLBEKKRRADALEEQRCISELK 740

Search completed: January 2, 2003, 12:36:40
Job time: 23 secs
